

University of Windsor

## Scholarship at UWindor

---

Electronic Theses and Dissertations

Theses, Dissertations, and Major Papers

---

8-30-2019

# Computational Insights into Bio-relevant Sulfur Chemistry

Sahar Nikoo

*University of Windsor*

Follow this and additional works at: <https://scholar.uwindsor.ca/etd>

---

### Recommended Citation

Nikoo, Sahar, "Computational Insights into Bio-relevant Sulfur Chemistry" (2019). *Electronic Theses and Dissertations*. 7827.

<https://scholar.uwindsor.ca/etd/7827>

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email ([scholarship@uwindsor.ca](mailto:scholarship@uwindsor.ca)) or by telephone at 519-253-3000ext. 3208.

# **Computational Insights into Bio-relevant Sulfur Chemistry**

By

**Sahar Nikoo**

A Dissertation

Submitted to the Faculty of Graduate Studies  
through the Department of Chemistry and Biochemistry  
in Partial Fulfillment of the Requirements for  
the Degree of Doctor of Philosophy  
at the University of Windsor

Windsor, Ontario, Canada

2019

© 2019 Sahar Nikoo

# Computational Insights into Bio-relevant Sulfur Chemistry

by

Sahar Nikoo

APPROVED BY:

---

N.M. Mora Diez, External Examiner  
Thompson Rivers University

---

L. Porter  
Biomedical Science

---

J. Green  
Department of Chemistry and Biochemistry

---

B. Mutus  
Department of Chemistry and Biochemistry

---

J. Gauld, Advisor  
Department of Chemistry and Biochemistry

August 30, 2019

## Declaration of Co-Authorship / Previous Publication

### I. Co-Authorship

I hereby declare that this dissertation incorporates material that is result of joint research, as follows:

*Chapter 3 of this dissertation was equally co-authored with Paul J. Meister under the supervision of professor James W. Gauld. In all cases, the key ideas, conducting computational calculation, data analysis, interpretation, and writing were equally performed by the author, and the co-author's contribution. The written manuscript was edited in a collaboration with Dr. John J. Hayward under the supervision of Prof. James W. Gauld.*

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my dissertation, and have obtained written permission from each of the co-author(s) to include the above material(s) in my dissertation.

I certify that, with the above qualification, this dissertation, and the research to which it refers, is the product of my own work.

### II. Previous Publications

This dissertation includes one original paper that has been previously published for publication in the peer-reviewed journal as follows:

Dissertation Chapter	Citation	Publication Status
Chapter 3	Nikoo, S.; Meister, P. J.; Hayward, J. J.; Gauld, J. W. <i>Molecules</i> <b>2018</b> , 23, 3323.	Published



I certify that I have obtained written permission from the copyright owners to include the above-published material in my dissertation. I certify that the above material describes work completed while registered as a graduate student at the University of Windsor.

### III. General

I declare that, to the best of my knowledge, my dissertation does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my dissertation, published or otherwise, are fully acknowledged in accordance with the standard referencing practices.

I declare that this is a true copy of my dissertation, including any final revisions, as approved by my dissertation committee and the Graduate Studies office, and that this dissertation has not been submitted for a higher degree to any other University or Institution.

## Abstract

Sulfur-containing molecules are chemically and functionally versatile compounds, exemplified by their diverse roles from enzymatic processes to organic synthesis and drug design. With the goal of gaining detailed and deeper insights into the chemistry of such species, multi-scale computational modeling techniques we have applied in this work.

**Chapter 1** provides a brief summary of the importance of sulfur, its functionality, and reactivity in biological systems including catalytic environments such as enzymes. In **Chapter 2**, an overview of the key features of the common and contemporary computational approaches is explained briefly.

In **Chapter 3**, systematic benchmark studies are performed to determine reliable and accurate structures as well as thermochemical data for a series of bio-relevant polysulfur/selenium-containing compounds. Of the variety of DFT functionals and Pople basis sets examined, the  $\omega$ B97XD/6-311G(2d,p) level of theory is found to generally give the most accurate and reliable results. Furthermore, S—S bond lengths are more sensitive to the choice of basis set than those containing Se. Comparison of the proton affinities and gas-phase basicities of thiols and their corresponding persulfide derivatives indicates that extending the sulfur chain decreases their values, suggesting that polysulfur species exist as deprotonated species in biological systems.

In **Chapter 4**, the roles of solvent choice on the possible mechanisms of formation of sulfonamides *via* the reaction of SO<sub>2</sub> and N-tosyl hydrazone using DFT-based methods in combination with implicit and hybrid implicit/explicit solvation models is examined. The results indicate that solvent-solute interactions can play critical roles in such reactions. Of the solvents considered, DMSO and piperidine are found to be the most effective (*i.e.*, actively involved) facilitating sulfonamide bond formation.

Applying DFT and conventional *ab initio* methods, **Chapter 5** examines the formation of SO<sub>2</sub>-containing molecules including sulfones, sulfonamides, and sulfamides *via* the

radical-based reaction of SO<sub>2</sub> with a systematic series of xiamycin-inspired aromatic C- and N-centered radicals. A preference for C–S(O<sub>2</sub>) vs. N–S(O<sub>2</sub>) bond formation is observed with formation of sulfones being thermodynamically preferred to sulfamides. Also, of the DFT functionals used, the M06-2X functional was shown to be most reliable for providing optimized geometries and relative energies of the SO<sub>2</sub>-containing species examined.

In **Chapter 6**, the formation of a range of possible HNO-derived post-translational modifications of cysteinyl and cysteinyl persulfide was examined using DFT-based methods. It is shown that the formation of the initial -X-NHOH (X=S, S-S) containing intermediate is independent of the residues position in the peptide while their subsequent reaction and final PTM formed is dependent on the residues position. More specifically, reaction of HNO with N-terminus or internal residues leads to formation of disulfide or sulfonamide (*e.g.*, Cys-SS-Cys or Cys-S(O)-NH<sub>2</sub>) *via* rearrangement and nucleophilic substitutions, respectively. Meanwhile, Cys-X-NH<sub>2</sub> derived from C-terminus peptide leads to Cys-X-OH formation through the intermediacy of a 5- or 6-membered cyclic intermediates in cystenyl and cystenyl persulfide, respectively.

In **Chapter 7** we examine the active site, substrate binding, and catalytic mechanism of the bacterial Ni(II)-dimethylsulfoniopropionate (DMSP) lyase (DddK) enzyme. The findings show that two active site tyrosyls (Tyr64 and 122) play significant roles in substrate binding, with Tyr64 also acting as a Lewis base to initiate the  $\beta$ -concerted elimination reaction to form the dimethyl sulfide product.

In **Chapter 8** we examine, using a multi-scale computational approach, a possible disulfide–sulfenylamide shuttling mechanism in the active site of DAH7PS enzyme. The results imply the key role of the metal ion (Mn(II)) and acidic environment in the potential interconversion between these conformations. Our findings infer that the preference of the cyclic sulfenylamide conformation to disulfide in the enzyme active site switches to the preference of disulfide to cyclic sulfenylamide conformation in the absence of metal ions and/or providing an acidic environment.

## **Dedication**

This dissertation is dedicated to my loving parents and family for their endless affection, sacrifices, and encouragement.

## **Acknowledgements**

I would like to take this opportunity to express my heartfelt gratitude to all of those who helped me to make my work a success. First and foremost, I would like to sincerely thank my esteemed supervisor, Professor James W. Gauld, for his sound advice and guidance, valuable suggestions, and endless patience and encouragement throughout this adventurous journey. Working under Professor Gauld's supervision has provided me with the opportunity to learn and gain invaluable skills and become a researcher.

I would like to extend my appreciation to my advisory committee members, Professor Bulent Mutus, Professor James Green, and Professor Lisa Porter for their fruitful discussions and meetings over the course of my graduate studies. Furthermore, I would like to thank them for taking their valuable time to assess this work. A special note of recognition and thanks goes to Professor Nelaine Mora-Diez for giving me the honor of being on my advisory committee as an external examiner.

I further extend my appreciation to all the past and present members of the Gauld Group that I had the pleasure to work and spend most of my time with throughout the last four years: Dr. Mohamed Aboelnga, Dr. Abayomi Faponle, Taqred Alnakhli, Travis Dewolfe, Wanlei Wei, and Anupom Roy. In particular, I would also like to express my gratitude to Dr. John Hayward and Paul Meister for their great collaboration on one of the studies presented in this work.

There are no words to express my profound gratitude and appreciation to my beloved parents and family members for their love and spiritual support during the long years of my academic journey. From the bottom of my heart, I am deeply thankful for having such wonderful people in my life, and this would not have been easy or even possible without their help and support.

And finally, I would like to acknowledge the Ontario Trillium Scholarship (OTS) and the Natural Sciences and Engineering Research Council of Canada (NSERC) for their

financial support and Compute Calcul Canada and SHARCNET for computational resources.

## Table of Contents

<b>Declaration of Co-Authorship / Previous Publication .....</b>	<b>iii</b>
<b>Abstract.....</b>	<b>v</b>
<b>Dedication .....</b>	<b>vii</b>
<b>Acknowledgements .....</b>	<b>viii</b>
<b>List of Tables .....</b>	<b>xvi</b>
<b>List of Figures.....</b>	<b>xviii</b>
<b>List of Schemes.....</b>	<b>xxv</b>
<b>List of Appendices.....</b>	<b>xxvii</b>
<b>List of Abbreviations .....</b>	<b>xxix</b>

### CHAPTER 1

#### Introduction

<b>1.1 Introduction.....</b>	<b>2</b>
<b>1.2 Biological Sulfur-Containing Compounds.....</b>	<b>2</b>
<b>1.3 Catalytic Solvents and Proteins .....</b>	<b>5</b>
<b>1.4 Computational Chemistry's Application .....</b>	<b>6</b>
<b>1.5 References .....</b>	<b>7</b>

### CHAPTER 2

#### Computational Methods

<b>2.1 Computational Chemistry.....</b>	<b>14</b>
<b>2.2 Molecular Mechanics.....</b>	<b>15</b>
<b>2.3 Molecular Dynamics (MD) Simulations.....</b>	<b>17</b>
<b>2.4 Quantum Mechanical (QM) Methods.....</b>	<b>18</b>
2.4.1 <i>Ab Initio</i> Based Methods .....	19
2.4.2 Density Functional Theory (DFT)-Based Methods .....	20
<b>2.5 Basis Set .....</b>	<b>24</b>

<b>2.6 Quantum Mechanical (QM)-Cluster Approach.....</b>	<b>25</b>
<b>2.7 Solvation Models .....</b>	<b>28</b>
<b>2.8 Quantum Mechanics/Molecular Mechanics (QM/MM).....</b>	<b>30</b>
2.8.1 QM and MM Interactions.....	31
2.8.2 Computing the Total Potential Energy of the QM/MM.....	33
2.8.3 QM/MM Boundary Schemes .....	34
<b>2.9 Thermodynamic Aspects .....</b>	<b>36</b>
<b>2.10 References .....</b>	<b>37</b>

## CHAPTER 3

### An Assessment of Computational Methods for Calculating Accurate Structures and Energies of Bio-Relevant Polysulfur/Selenium-Containing Compounds

<b>3.1 Introduction.....</b>	<b>46</b>
<b>3.2 Computational Methods.....</b>	<b>48</b>
<b>3.3 Results and Discussion.....</b>	<b>50</b>
3.3.1 Structural Assessment of CH <sub>3</sub> XXH and CH <sub>3</sub> XX <sup>-</sup> (X=S, Se) .....	50
3.3.2.1 Method Sensitivity to Basis Set Changes .....	50
3.3.2.2 Effects of Increasing Basis Set Size .....	53
3.3.2 Effect of Conjugation: CH <sub>2</sub> =CHXXH and CH <sub>2</sub> =CHXX <sup>-</sup> (X=S, Se).....	55
3.3.3 Effects of Mixed Sulfur/Selenium: CH <sub>3</sub> XYH (X=S, Se; Y=Se, S) .....	56
3.3.4 Extending a Sulfide: CH <sub>3</sub> SSSH and CH <sub>2</sub> CHSSSH.....	57
3.3.5 Obtaining Reliable and Accurate Thermochemistry for CH <sub>3</sub> SSH and CH <sub>2</sub> CHSSH .....	59
3.3.6 Thermochemistry of Selenium-Containing Species.....	62
3.3.7 Extending the Persulfides to Trisulfides .....	65
3.3.8 Cysteine-Derived Polysulfides: CysSS <sub>n</sub> H ( <i>n</i> =1-3).....	66
<b>3.4 Conclusions.....</b>	<b>69</b>
<b>3.5 References .....</b>	<b>71</b>



## CHAPTER 4

### Computational Investigations Into Implicit and Explicit Roles of Solvent In Sulfonamide Formation

<b>4.1 Introduction.....</b>	<b>78</b>
<b>4.2 Computational Methods.....</b>	<b>81</b>
<b>4.3 Results and Discussion.....</b>	<b>81</b>
4.3.1 Reactivity of SO <sub>2</sub> toward N-tosyl Hydrazone .....	82
4.3.1.1 Implicit Solvent Model.....	82
4.3.1.2 Hybrid Explicit and Implicit Solvent Model .....	84
4.3.2 Reactivity of SO <sub>2</sub> toward Diaza Species .....	93
4.3.3 Reactivity of Sulfene toward Amines .....	94
4.3.4 Alternative Sulfonamide Formation from N-tosyl Hydrazone .....	97
<b>4.4 Conclusions.....</b>	<b>98</b>
<b>4.5 References .....</b>	<b>100</b>

## CHAPTER 5

### Computational Insights Into the Non-Enzymatic XiaH Facilitated Radical-Based Formation of SO<sub>2</sub>-Containing Compounds

<b>5.1 Introduction.....</b>	<b>107</b>
<b>5.2 Computational Methods.....</b>	<b>109</b>
<b>5.3 Results and Discussion.....</b>	<b>110</b>
5.3.1 Pyrrolyl.....	111
5.3.1.1 N–S(O <sub>2</sub> ) Bond Formation.....	111
5.3.1.2 C–S(O <sub>2</sub> ) Bond Formation .....	112
5.3.1.3 Cross-Linking Reactions .....	113
5.3.2. Indolyl .....	114
5.3.2.1 N–S(O <sub>2</sub> ) Bond Formation.....	114
5.3.2.2 C–S(O <sub>2</sub> ) Bond Formation .....	115

5.3.2.3 Cross-Linking Reactions .....	116
5.3.3 Cabazolyl.....	117
5.3.3.1 N–S(O <sub>2</sub> ) Bond Formation.....	117
5.3.3.2 C–S(O <sub>2</sub> ) Bond Formation .....	118
5.3.3.3 Cross-Linking Reactions .....	119
5.3.3.4 Substitution Effects.....	120
5.3.3.5 Cross-Linking Reactions .....	121
5.3.4 Radical Xiamycin.....	122
5.3.4.1 N–S(O <sub>2</sub> ) Bond Formation.....	123
5.3.4.2 C–S(O <sub>2</sub> ) Bond Formation .....	124
5.3.4.3 Cross-Linking Reactions .....	124
<b>5.4 Conclusions.....</b>	<b>126</b>
<b>5.5 References .....</b>	<b>127</b>

## CHAPTER 6

### Computational Study on HNO-Derived Post-Translational Modifications of Cysteinyl and Cysteinyl Persulfide

<b>6.1 Introduction.....</b>	<b>135</b>
<b>6.2 Computational Methods.....</b>	<b>138</b>
<b>6.3 Results and Discussion.....</b>	<b>139</b>
6.3.1 Reaction of HNO and Cysteinyl Residue.....	139
6.3.1.1 N-hydroxysulfenamide Formation .....	139
6.3.1.2 Reactions Resulted of Protonated N-Hydroxysulfenamide Derived from Internal Cysteinyl and HNO .....	145
6.3.1.3 Reactions Resulted of Protonated N-Hydroxysulfenamide Derived from C- terminus Cysteinyl and HNO .....	148
6.3.2 Reaction of HNO and Cysteinyl Persulfide .....	150
6.3.2.1 Cys-SS-NHOH Formation.....	150

6.3.2.2 Reactions Resulted of Cys-SS-NH <sub>2</sub> <sup>+</sup> OH Derived from Internal Cysteiny	
Persulfide and HNO.....	152
6.3.2.3 Reactions Resulted of Cys-SS-NH <sub>2</sub> <sup>+</sup> -OH Derived from C-terminus Cysteiny	
Persulfide and HNO.....	153
6.3.3 Cys-S-NHOH <i>versus</i> Cys-S-ONH <sub>2</sub> .....	154
6.3.4 Reactivity of HNO toward Cysteine Sulfenic Acid .....	155
<b>6.4 Conclusions.....</b>	<b>157</b>
<b>6.5 References .....</b>	<b>159</b>

## CHAPTER 7

### Computational Study on the Mechanism of Bacterial Ni(II)- Dimethylsulfoniopropionate Lyase

<b>7.1 Introduction.....</b>	<b>166</b>
<b>7.2 Computational Methods.....</b>	<b>169</b>
7.2.1 Molecular Dynamics (MD) Simulations .....	169
7.2.2 QM-Cluster and QM/MM Calculations .....	170
<b>7.3 Results and Discussion.....</b>	<b>171</b>
7.3.1 Ni-DddK Enzyme.....	171
7.3.1.1 QM-Cluster Investigation .....	171
7.3.1.2 QM/MM Investigation.....	175
7.3.2 Mn-DddK Enzyme .....	179
<b>7.4 Conclusion .....</b>	<b>182</b>
<b>7.5 References .....</b>	<b>183</b>

## CHAPTER 8

### Computational Insights into A Potential Disulfide-Sulfenylamide Exchange Mechanism in 3-Deoxy-d-Arabino-Heptulosonate 7-Phosphate Synthase (DAH7PS)

<b>8.1 Introduction.....</b>	<b>191</b>
<b>8.2 Computational Methods.....</b>	<b>194</b>

8.2.1 Molecular Dynamics (MD) Simulations .....	194
8.2.2 QM-Cluster and QM/MM Calculations .....	195
<b>8.3 Results and Discussion.....</b>	<b>197</b>
8.3.1 QM-Cluster Investigation.....	197
8.3.2 QM/MM Investigation; Small Model of QM Region .....	199
8.3.2.1 Mn(II) Influence on the Exchange Mechanism .....	200
8.3.2.2 Proton Transferring from Threonine .....	202
8.3.3 QM/MM Investigation; Large Model of QM Region .....	204
8.3.3.1 Mn(II) Influence on the Exchange Mechanism .....	207
8.3.4 Molecular Dynamics (MD) Simulations .....	208
<b>8.4 Conclusion .....</b>	<b>211</b>
<b>8.5 References .....</b>	<b>214</b>

## CHAPTER 9

### Conclusions and Future Work

<b>9.1 Conclusions.....</b>	<b>221</b>
<b>APPENDICES .....</b>	<b>227</b>
<b>Appendix A .....</b>	<b>227</b>
<b>Appendix B .....</b>	<b>237</b>
<b>VITA AUCTORIS .....</b>	<b>239</b>

## List of Tables

<b>Table 3.1.</b> Selected optimized bond lengths in ångström (Å) for CH <sub>3</sub> XXH and CH <sub>3</sub> XX <sup>-</sup> (X=S, Se). .....	52
<b>Table 3.2.</b> Selected optimized bond lengths (ångström, Å) for CH <sub>2</sub> =CHXXH and CH <sub>2</sub> =CHXX <sup>-</sup> (X=S, Se). .....	55
<b>Table 3.3.</b> Selected optimized bond lengths (ångström, Å) for CH <sub>3</sub> XYH and CH <sub>3</sub> XY <sup>-</sup> (X=S, Se; Y=Se, S). .....	57
<b>Table 3.4.</b> Selected optimized bond lengths (Å) for RSSSH (R=CH <sub>3</sub> , CH <sub>2</sub> =CH). .....	58
<b>Table 3.5.</b> Homolytic S–S bond dissociation enthalpy (BDE) of RSSH, proton affinity (PA) and gas-phase basicity (GPB) of RSS <sup>-</sup> , and hydrogen affinity (HA) of RSS <sup>•</sup> (R=CH <sub>3</sub> , CH <sub>2</sub> CH). All energies calculated at 298.15K and in kJ mol <sup>-1</sup> . .....	61
<b>Table 3.6.</b> Calculated homolytic Se–Se bond dissociation enthalpies (BDE) of RSeSeH, proton affinity (PA) and gas-phase basicities (GPB) of RSeSe <sup>-</sup> , and Se–H homolytic bond dissociation enthalpies (HA) of RSeSe <sup>•</sup> (R=CH <sub>3</sub> , CH <sub>2</sub> CH). All energies in kJ mol <sup>-1</sup> . .....	64
<b>Table 3.7.</b> Homolytic RS–SH bond dissociation enthalpy (BDE) of RSSSH, proton affinity (PA) and gas-phase basicity (GPB) of RSSS <sup>-</sup> , and hydrogen affinity (HA) of RSSS <sup>•</sup> (R=CH <sub>3</sub> , CH <sub>2</sub> CH). All energies calculated at 298.15K and in kJ mol <sup>-1</sup> . .....	66
<b>Table 3.8.</b> Selected optimized bond lengths (ångström, Å) obtained at the ωB97XD/6-311G(2d,p) level of theory for C <sub>ys</sub> SS <sub>n</sub> H (n=0-3) and C <sub>ys</sub> SS <sub>n</sub> <sup>-</sup> (n=0-3). .....	67
<b>Table 5.1.</b> Calculated formation energy (kJ/mol) of SO <sub>2</sub> -containing compounds using pyrrolyl. ....	113
<b>Table 5.2.</b> Calculated formation energy (kJ/mol) of SO <sub>2</sub> -containing compounds using indolyl. ....	116
<b>Table 5.3.</b> Calculated formation energy (kJ/mol) of SO <sub>2</sub> -containing compounds applying carbazolyl. ....	120

<b>Table 5.4.</b> Calculated formation energy (kJ/mol) of SO <sub>2</sub> -containing compounds from substituted carbazolyl.....	123
<b>Table 5.5.</b> Calculated formation energy (kJ/mol) of SO <sub>2</sub> -containing compounds from xiamycin.....	125

## List of Figures

<b>Figure 1.1.</b> Examples of biologically relevant cysteine functional modification intermediates and products. <sup>7</sup> .....	3
<b>Figure 1.2.</b> The protective role of S—N bond formation due to oxidative stress in biological systems. <sup>28</sup> .....	5
<b>Figure 2.1.</b> Number of publications per year; a result of the search on ISI Web of Science using the topic of “Computational Chemistry” (accessed on 28 <sup>th</sup> March 2019). .....	15
<b>Figure 2.2.</b> Representation of the bonded and non-bonded interactions calculated by molecular mechanics.....	16
<b>Figure 2.3.</b> Schematic representation of correlation energy calculation.....	19
<b>Figure 2.4.</b> Schematic representation of the Perdew's ‘Jacob ladder’ for 5-generated DFT functionals. <sup>22</sup> .....	21
<b>Figure 2.5.</b> Schematic representation of Pople basis sets compositions. ....	25
<b>Figure 2.6.</b> Schematic illustration of the explicit solvation model; solute with some solvent molecules included in the calculation. ....	29
<b>Figure 2.7.</b> Schematic illustration of Tomasi's PCM solvation model. ....	30
<b>Figure 2.8.</b> Schematic representation of a QM/MM model. The shaded and center regions correspond to the MM and QM layers, respectively.....	31
<b>Figure 2.9.</b> Schematic representation of an additive coupling within a QM/MM methodology in computing total energy. ....	33
<b>Figure 2.10.</b> Schematic representation of a subtractive coupling within a QM/MM methodology in computing total energy. ....	34
<b>Figure 2.11.</b> Illustration of different procedures to cap the QM layer. ....	36
<b>Figure 3.1.</b> Optimized structures (with C—S and S—S bond length given in ångström (Å)) obtained at the $\omega$ B97XD/6-311G(2d,p) level of theory for neutral cysteine and its	

polysulfides ( $\text{CysSS}_n\text{H}$ ,  $n=1-3$ ) derivatives (top row), and their corresponding anionic deprotonated derivatives ( $\text{CysSS}_n^-$ ,  $n=0-3$ ; bottom row)..... 67

**Figure 3.2.** Calculated (at the  $\omega\text{B97XD/6-311G(2d,p)}$  level of theory) values of **(A)**  $\text{RS}_n\text{—SH}$  BDE's for  $\text{CysSS}_n\text{H}$  ( $n=1-3$ ), **(B)** PA's (blue) and GPB's (orange) of  $\text{CysSS}_n^-$  ( $n=0-3$ ), and **(C)** HA's of  $\text{CysSS}_n^\bullet$  ( $n=0-3$ ). All energies are in  $\text{kJ mol}^{-1}$ ..... 69

**Figure 4.1.** Calculated (electronic energy) activation barriers (difference in energy between reactants and transition structure; blue line) and reaction energies (difference in energy between reactants and products; green line) for the reaction of  $\text{SO}_2$  with N-tosyl hydrazone in various solvents and in the gas-phase. .... 83

**Figure 4.2.** Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of  $\text{SO}_2$  toward N-tosyl hydrazone in the presence of different bases with the calculated proton affinity (PA) of each applied base ( $\text{kJ/mol}$ ) in DMSO..... 89

**Figure 4.3.** Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of  $\text{SO}_2$  toward N-tosyl hydrazone in the presence of water and DMSO as the base with the calculated proton affinity (PA) of each applied base ( $\text{kJ/mol}$ ) in toluene. . 90

**Figure 4.4.** Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of  $\text{SO}_2$  toward N-tosyl hydrazone in toluene, DMSO, and water with applying **A) DMSO B) water** as the base..... 91

**Figure 4.5.** Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of  $\text{SO}_2$  toward N-tosyl hydrazone in an environment with the same implicit and explicit solvent molecule with the calculated proton affinity (PA) in  $\text{kJ/mol}$ . .... 91

**Figure 4.6.** Potential energy surface for the formation of  $\beta,\gamma$ -unsaturated sulfinic acid and its decomposition to sulfene in DMSO..... 92

**Figure 4.7.** Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of  $\text{SO}_2$  toward diaza species in various solvents and in the gas-phase. .... 93

**Figure 4.8.** Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of sulfene toward piperidine in various solvents and in the gas-phase..... 94



<b>Figure 4.9.</b> Potential energy surface for the formation of sulfonamide from sulfene in the presence of two equivalents of piperidine in DMSO.....	96
<b>Figure 4.10.</b> Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of sulfene toward piperidine with/without the presence of different bases in DMSO.....	97
<b>Figure 5.1.</b> Spin density map for indole sulfonyl species using (A) B3LYP and (B) M06-2X functionals, respectively.....	114
<b>Figure 5.2.</b> Spin density map for carbazole sulfonyl species using (A) B3LYP and (B) M06-2X functionals, respectively.....	118
<b>Figure 6.1.</b> Optimized reactant complexes (RC <sub>1</sub> ), transition structures (TS <sub>1</sub> ), and product complexes (PC <sub>1</sub> ) with selected bond distances in ångströms (Å) of the water-catalyzed reaction of HNO with C-terminus, internal, and N-terminus cysteine residues from up to down, respectively. ....	140
<b>Figure 6.2.</b> QTAIM molecular graphs of the obtained transition structures of the water-catalyzed reaction of HNO with C-terminus, internal, and N-terminus cysteine residues from left to right, respectively. ....	141
<b>Figure 6.3.</b> Optimized transition structures with selected bond distances in ångströms (Å) and QTAIM molecular graphs of the reaction of HNO with C-terminus, internal, and N-terminus cysteine residues from left to right, respectively. ....	142
<b>Figure 6.4.</b> Optimized 6-membered cyclic transition structures in $\epsilon=78.35$ and $\epsilon=4$ with selected bond distances shown in ångströms (Å) for the water-catalyzed reaction of HNO with N-terminus cysteine residue.....	143
<b>Figure 6.5.</b> Obtained potential energy surface for the formation of N-hydroxysulfenamide. The green, black and blue colored surface is for the reactivity of HNO toward N-terminus, internal, and C-terminus cysteine residues, respectively. Values in parentheses and brackets are calculated in $\epsilon=10$ and 4, respectively.....	144

<b>Figure 6.6.</b> Potential energy surface with optimized structures and selected bond distances in ångstroms (Å) for the formation of sulfinamide.H <sup>+</sup> . The black and blue colored surfaces are for the water-catalyzed and uncatalyzed reaction. ....	147
<b>Figure 6.7.</b> Potential energy surface of the water-catalyzed sulfenic acid formation with optimized structures and selected bond lengths in ångstroms (Å). ....	149
<b>Figure 6.8.</b> Optimized transition structures with selected bond distances in ångstroms (Å) and QTAIM molecular graphs of the water-catalyzed reaction of HNO with C-terminus, internal, and N-terminus cysteine persulfide residues. ....	150
<b>Figure 6.9.</b> Obtained potential energy surface for the formation of Cys-SS-NH-OH. The green, black and blue colored surface is for the reactivity of HNO toward N-terminus, internal, and C-terminus cysteinyl persulfide, respectively. ....	151
<b>Figure 6.10.</b> Potential energy surface with optimized structures and selected bond distances in ångstroms (Å) for the formation of Cys-SS <sup>+</sup> (OH)-NH <sub>2</sub> . The black and blue colored surface is for a water-assisted and unassisted catalyzed reaction. ....	153
<b>Figure 6.11.</b> Potential energy surface of the water-catalyzed Cy-SS-OH formation with optimized structures and selected bond lengths in ångstroms (Å). ....	154
<b>Figure 6.12.</b> Potential energy surface of reactivity of O and N atoms in HNO toward cysteine residue (green colored surface) and cysteine sulfenic acid (blue colored surface) with the schematic representation of transition structures. ....	156
<b>Figure 7.1.</b> Schematic illustration of the optimized structures of the unbound and bound active sites using the QM-cluster approach with the selected bond distances in ångstroms (Å). ....	172
<b>Figure 7.2.</b> Schematic illustration of the optimized structures of the reactive complex, transition structure, and product complex using the QM-cluster approach with selected bond distances in ångstroms (Å). ....	174

<b>Figure 7.3.</b> Free energy surface obtained using the QM-cluster approach; (A) M06 functional (B) M06L functionals. The green, black, and blue colored surfaces are for the 6-31G(d), 6-31G(d,p), and 6-311G(d,p) level of optimization, respectively. ....	175
<b>Figure 7.4.</b> Schematic representation of the optimized unbound and bound active site complexes using the ONIOM(QM/MM) approach, with selected bond distances in ångstroms (Å).....	176
<b>Figure 7.5.</b> Schematic drawing of the optimized molecular optimized structures of the reactive complex, transition structure, and product complex using the ONIOM(QM/MM) approach with selected bond distances in ångstroms (Å). ....	177
<b>Figure 7.6.</b> Calculated free energy surfaces using the ONIOM(QM/MM) approach; (A) M06 (B) M06L functionals, respectively. The green, black, and blue colored surfaces are for 6-31G(d), 6-31G(d,p), and 6-311G(d,p) level of optimization, respectively. Values in parentheses are calculated applying the electrostatic formalism on the single-point energy. ....	179
<b>Figure 7.7.</b> Calculated free energy surfaces using the QM-cluster approach in Mn-DddK enzyme. The black and blue colored surfaces are for M06/6-31G(d,p) and M06L/6-31G(d,p) level of optimization, respectively. ....	180
<b>Figure 7.8.</b> QM/MM optimized structures of <b>RC</b> <sub>4</sub> , <b>TS</b> <sub>4</sub> , and <b>PC</b> <sub>4</sub> with selected bond distances in ångstroms (Å) and the relative free energy of <b>TS</b> <sub>4</sub> and <b>PC</b> <sub>4</sub> with respect to the <b>RC</b> <sub>4</sub> in Mn-centered coordination. Black and blue colored values are related to M06 and M06L functionals, respectively. Values in parentheses are calculated utilizing the electrostatic formalism on the single point energy. ....	181
<b>Figure 8.1.</b> Illustration of two different active sites of DAH7P enzyme (PDB ID: 3NUE) with cyclic sulfenylamide (on the left) and disulfide bond (on the right) and selected key distances in ångstroms (Å).....	193

<b>Figure 8.2.</b> Schematic representation of optimized structures of the active site ( <b>A</b> ) and protonated oxygen amide in the active site ( <b>B</b> ) with the selected bond distances shown in ångstroms (Å) at the B3LYP/6-31G(d,p) level of theory. ....	198
<b>Figure 8.3.</b> Schematic illustration of optimized structures of the active site ( <b>C</b> ) and protonated oxygen amide in the active site ( <b>D</b> ), and a complex result of examining proton transferring from Tyr436 to oxygen atom of sulfenylamide bond ( <b>E</b> ) with the selected distances in ångstroms (Å) using ONIOM(B3LYP/6-31G(d,p):AMBER96). ....	201
<b>Figure 8.4.</b> Schematic representation of optimized complexes of the active site ( <b>F</b> ) and the result of proton transferred from Tyr to oxygen amide in the active site ( <b>G</b> ) without metal ion with selected distances shown in ångstroms (Å) using ONIOM(B3LYP/6-31G(d,p):AMBER96). ....	202
<b>Figure 8.5.</b> Schematic representation of optimized complexes of the active sites including Thr438 ( <b>H</b> ) and the complex resulted from proton transferring from Thr438 to oxygen amide ( <b>I</b> ) with the selected distances in ångstroms (Å) using ONIOM(B3LYP/6-31G(d,p):AMBER96). ....	203
<b>Figure 8.6.</b> Schematic illustration of the optimized complexes including disulfide result of the hydronium ion presence ( <b>J</b> ), and proton transferring from Thr436 to oxygen amide in the presence of hydronium ion ( <b>K</b> ) with the selected distances in ångstroms (Å) using ONIOM(B3LYP/6-31G(d,p):AMBER96). ....	204
<b>Figure 8.7.</b> Schematic representation of optimized structures of the active site ( <b>L</b> ) and protonated oxygen amide in the active site ( <b>M</b> ) and a structure result of examining proton transferring from Tyr436 to oxygen atom of sulfenylamide bond ( <b>N</b> ) with the larger model of QM region using ONIOM(B3LYP/6-31G(d,p):AMBER96) with the selected distances in ångstroms (Å). ....	206
<b>Figure 8.8.</b> Schematic illustration of optimized complexes of the active site ( <b>O</b> ) and the result of proton transferred from Tyr436 to oxygen amide in the active site ( <b>P</b> ) with the	

larger model of QM region in the absence of metal ion with selected distances shown in ångstroms (Å) using ONIOM(B3LYP/6-31G(d,p):AMBER96). .....	208
<b>Figure 8.9.</b> Plot of the calculated RMSD of QM layer and its surrounding ( <b>A</b> ) and distance change between $C_{87}S$ and Mn(II) with respect to the time ( <b>B</b> ) in a system including cyclic sulfenylamide. ....	210
<b>Figure 8.10.</b> RMSD of the QM layer and its surrounding in a system including disulfide. .....	211
<b>Figure 8.11.</b> Plot of the second active site and its surrounding RMSDs in the enzyme including disulfide. ....	211

## List of Schemes

<b>Scheme 1.1.</b> Some natural products containing varied sulfur-based functional groups. <sup>5-6</sup>	2
<b>Scheme 3.1.</b> Several examples of naturally occurring polysulfur species. <sup>17, 28</sup>	47
<b>Scheme 3.2.</b> Schematic illustration of the species considered in this study.	49
<b>Scheme 4.1.</b> Schematic representative of selected natural and therapeutic sulfonamides species. <sup>11, 21</sup>	79
<b>Scheme 4.2.</b> The ene reaction of SO <sub>2</sub> and hydrazones.	82
<b>Scheme 4.3.</b> Schematic illustration of the optimized structures with the selected bond lengths in ångströms (Å) for the reaction of SO <sub>2</sub> and N-tosyl hydrazone in DMSO.	82
<b>Scheme 4.4.</b> Schematic illustration of the optimized structures with the selected bond lengths in ångströms (Å) for the reaction of SO <sub>2</sub> and N-tosyl hydrazone in the DMSO-based system.	85
<b>Scheme 4.5.</b> Schematic illustration of the optimized structures with the selected bond lengths in ångströms (Å) for the reaction of SO <sub>2</sub> and N-tosyl hydrazone in the HN <sub>3</sub> -based system.	86
<b>Scheme 4.6.</b> Schematic illustration of the optimized structures with the selected bond lengths in ångströms (Å) for the reaction of SO <sub>2</sub> and N-tosyl hydrazone in the piperidine-based system.	87
<b>Scheme 4.7.</b> Schematic illustration of the optimized structures with the selected bond lengths in ångströms (Å) for the reaction of SO <sub>2</sub> and N-tosyl hydrazone in the water-based system.	88
<b>Scheme 4.8.</b> Decomposition reaction of $\beta,\gamma$ -unsaturated sulfinic acid to sulfene.	92
<b>Scheme 4.9.</b> Reactivity of diazo species toward SO <sub>2</sub> to form sulfene.	93
<b>Scheme 4.10.</b> Reactivity of sulfene toward piperidine to form sulfonamide.	94
<b>Scheme 4.11.</b> Schematic description of the obtained transition structures with the selected bond lengths in ångströms (Å) of solvent-assisted addition reaction of piperidine to sulfene	

in DMSO with the active role of the water molecule (TS <sub>7</sub> ), the passive role of the water molecule (TS <sub>8</sub> ), and the passive role of the DMSO (TS <sub>9</sub> ).....	95
<b>Scheme 4.12.</b> Formation of sulfonamide from the reaction of sulfene and two equivalents of piperidine. ....	96
<b>Scheme 4.13.</b> Plausible products of the reaction between (piperidine–SO <sub>2</sub> ) <sup>–</sup> complex and CH <sub>3</sub> CH <sup>+</sup> Ph in DMSO with corresponding formation energy. ....	98
<b>Scheme 5.1.</b> Some examples of biologically active sulfone and sulfonamides. <sup>19</sup> .....	108
<b>Scheme 5.2.</b> Schematic initial illustration of the radical species considered in this study. ....	110
<b>Scheme 6.1.</b> Schematic illustration of the reaction and resulting products of HNO with a: (A) thiol, and (B) C-terminus cysteinyl thiol.....	136
<b>Scheme 6.2.</b> Schematic illustration of the chemical models used for the cysteinyl (X = S) and cysteinyl persulfide (X = S-S) residues within (internal) and in the C- and N-terminus positions, of a peptide. ....	138
<b>Scheme 6.3.</b> Schematic representation of N-hydroxysulfenamide formation.....	139
<b>Scheme 6.4.</b> Schematic representation of reaction of N-hydroxysulfenamide to form sulfinamide and disulfide.....	145
<b>Scheme 6.5.</b> Schematic illustration of cysteine sulfenic acid formation from N-hydroxysulfenamide.....	149
<b>Scheme 6.6.</b> Schematic representation of Cys-SS <sup>+</sup> (OH)-NH <sub>2</sub> formation from Cys-SS-NH <sub>2</sub> <sup>+</sup> OH. ....	152
<b>Scheme 6.7.</b> Schematic representation of Cys-SS-OH formation from Cys-SS-NHOH. ....	154
<b>Scheme 7.1.</b> Schematic representation of the proposed Ni-DddK catalysed decomposition of DMSP. <sup>32</sup> .....	168

## List of Appendices

<b>Appendix A3.1.</b> Selected optimized bond lengths in ångström (Å) for CH <sub>3</sub> XXH and CH <sub>3</sub> XX <sup>-</sup> (X=S, Se) using MP2.....	228
<b>Appendix A3.2.</b> C–S and S–S bond lengths (ångström, Å) for CH <sub>2</sub> CHSSH and CH <sub>2</sub> CHSS <sup>-</sup> .....	229
<b>Appendix A3.3.</b> C–Se and Se–Se bond lengths (ångström, Å) for CH <sub>2</sub> CHSeSeH and CH <sub>2</sub> CHSeSe <sup>-</sup> .....	230
<b>Appendix A3.4.</b> C–Se, C–S and S–Se bond lengths (ångström, Å) for CH <sub>3</sub> XYH and CH <sub>3</sub> XY <sup>-</sup> (X=S, Se; Y=Se, S).....	231
<b>Appendix A3.5.</b> Optimized S–S bond lengths (ångström, Å) for RSSSH/ <sup>-</sup> (R=CH <sub>3</sub> , CH <sub>2</sub> =CH). .....	232
<b>Appendix A3.6.</b> Mulliken charges on every S of RSSSH/ <sup>-</sup> (R=CH <sub>3</sub> , CH <sub>2</sub> =CH).....	233
<b>Appendix A3.7.</b> Homolytic S–S bond dissociation enthalpy (BDE) of RSSH, proton affinity (PA) and gas-phase basicity (GPB) of RSS <sup>-</sup> , and hydrogen affinity (HA) of RSS <sup>•</sup> (R=CH <sub>3</sub> , CH <sub>2</sub> CH). All energies calculated at 298.15K and in kJ mol <sup>-1</sup> .....	233
<b>Appendix A3.8.</b> Homolytic Se–Se bond dissociation enthalpy (BDE) of RSeSeH, proton affinity (PA) and gas-phase basicity (GPB) of RSeSe <sup>-</sup> , and hydrogen affinity (HA) of RSeSe <sup>•</sup> (R=CH <sub>3</sub> , CH <sub>2</sub> CH). All energies calculated at 298.15K and in kJ mol <sup>-1</sup> .....	234
<b>Appendix A3.9.</b> Homolytic X–Y bond dissociation enthalpy (BDE) of RXYH, proton affinity (PA) and gas-phase basicity (GPB) of RXY <sup>-</sup> , and hydrogen affinity (HA) of RXY <sup>•</sup> (R=CH <sub>3</sub> , CH <sub>2</sub> CH; X=S, Se; Y=Se, S). All energies calculated at 298.15K and in kJ mol <sup>-1</sup> . .....	235
<b>Appendix A3.10.</b> Homolytic S–S bond dissociation enthalpy (BDE) of RSSSH, proton affinity (PA) and gas-phase basicity (GPB) of RSSS <sup>-</sup> , and hydrogen affinity (HA) of RSSS <sup>•</sup> (R=CH <sub>3</sub> , CH <sub>2</sub> CH). All energies calculated at 298.15K and in kJ mol <sup>-1</sup> .....	236



**Appendix B4.1.** Calculated Gibbs's Free Energy activation barriers (difference in energy between reactants and transition structure; blue line) and Gibbs's Free Energis of reactions (difference in energy between reactants and products; green line) for the reaction of SO<sub>2</sub> with N-tosyl hydrazone in various solvents and in the gas-phase. It is noted that the solvents are shown in order of decreasing dielectric constant. .... 238

**\*Appendix C** includes all the Gaussian archive entries of all optimized geometries obtained for chapters 3-8. To access Appendix C contact Prof. James Gauld at [gauld@uwindsor.ca](mailto:gauld@uwindsor.ca).

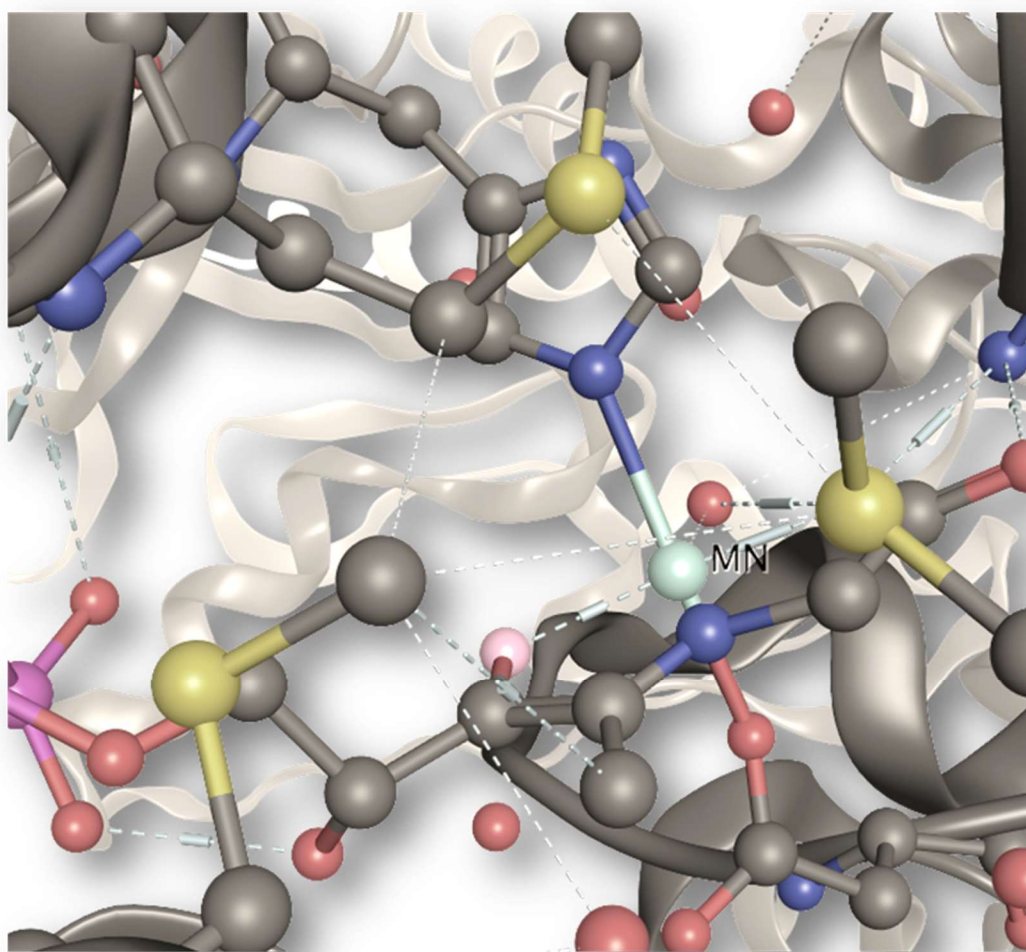
### **List of Abbreviations**

AMBER	Assisted Model Building with Energy Refinement
B3	Becke's 3-Parameters Exchange Functional
BCP	Bond Critical Point
BDE	Bond Disassociation Energy
CHARMM	Chemistry at Harvard Macromolecular Mechanics
CI	Configuration Interaction
CID	Configuration Interaction, Doubles
CISD	Configuration Interaction, Singles and Doubles
DFT	Density Functional Theory
DMS	Dimethyl Sulfide
DMSO	Dimethyl Sulfoxide
DMSP	Dimethyl Sulfoniopropionate
EE	Electronic Embedding
FCI	Full Configuration Interaction
GHO	Generalized Hybrid Orbital
GPB	Gas Phase Basicity
GROMOS	Groningen Molecular Simulation
GROMACS	Groningen Machine for Chemical Simulation
HA	Hydrogen Affinity
HF	Hartree-Fock
IC	Intermediate Complex
IEFPCM	Integral Equation Formalism Polarizable Continuum Model
IMOMM	Integrated Molecular Orbital Molecular Mechanics
IMOMO	Integrated Molecular Orbital Molecular Orbital
LSCF	Localized Self-Consistent Field
LYP	Lee-Yang-Parr Correction Functional

MD	Molecular Dynamics
ME	Mechanical Embedding
MM	Molecular Mechanics
MOE	Molecular Operating Environment
NAMD	Nanoscale Molecular Dynamics
ONIOM	Our Own N-Layered Integrated Molecular Orbital and Molecular Mechanics
PA	Proton Affinity
PC	Product Complex
PCM	Polarizable Continuum Model
PDB	Protein Data Bank
PES	Potential Energy Surface
PTM	Post-Translational Modification
QCI	Quadratic Configuration Interaction
QCISD	Quadratic Configuration Interaction, Singles and Doubles
QM	Quantum Mechanics
QM/MM	Quantum Mechanics/Molecular Mechanics
QTAIM	Quantum Theory of Atoms in Molecules
RC	Reactant Complex
RMSD	Root Mean Square Deviation
SCF	Self-Consistent Field
TS	Transition State

## CHAPTER 1

### Introduction

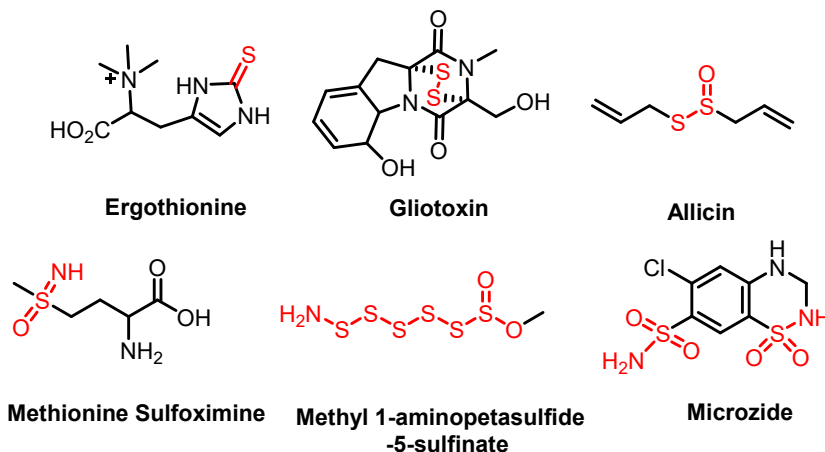


## 1.1 Introduction

Sulfur-based functional groups are of great importance to living organisms, biomolecules, natural products, drug design and development, as well as many industrial synthetic processes. Indeed, the importance of sulfur compounds originates from their unique chemistry and impressive capabilities in biochemical extension and other widespread applications within industrial and biochemical contexts.<sup>1</sup>

In general, sulfur's electronic configuration, its high polarizability, the availability of *d*-orbitals for bonding, and its low oxidation/reduction potential permits sulfur bond expansion from divalent compounds to valencies of 4 and 6.<sup>2-4</sup> Sulfur can also reach oxidation states ranging from -2 to +6, which increases the diversity of sulfur-containing functional groups in (bio)chemical systems.<sup>2-4</sup> **Scheme 1.1** highlights some natural products that contain sulfur-based functional groups in their molecular structures.

**Scheme 1.1.** Some natural products containing varied sulfur-based functional groups.<sup>5-6</sup>



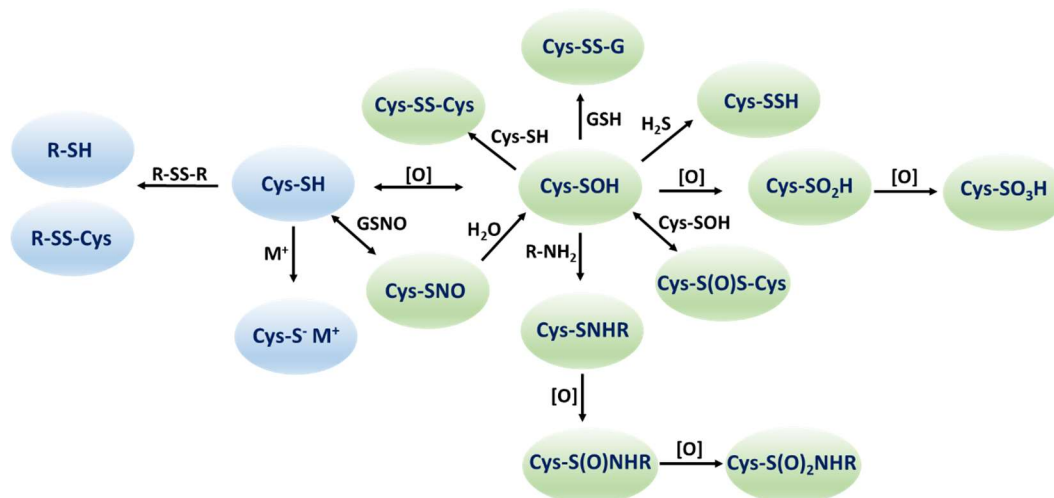
## 1.2 Biological Sulfur-Containing Compounds

In biomolecular systems, sulfur has a unique identity in proteins as the constituent of the sulfur-containing amino acids cysteine (Cys) and methionine (Met). As the least abundant amino acid residue (1–2%), cysteine plays substantial roles in protein stability, protein

function regulation, signal transduction, and binding to transition metal ions such as Zn, Cu, Fe, and is also involved in enzyme catalysis.<sup>7-8</sup>

The presence of the sulfur atom, with its diverse and rich chemistry allows a wide range of chemical reactivity that makes cysteine attractive to several post-translational modifications (PTMs) in biological systems among other amino acids.<sup>2, 8-11</sup> Post-translational modification is a powerful approach to alter and amplify protein function, regulation, and change the proteins' biological and therapeutic properties.<sup>12-14</sup> Today, identification of post-translational modification of biomolecules is a rapidly growing field, revealing potential for new reactions and applications.<sup>2, 12-14</sup>

In general, oxidative post-translational modification of cysteine results in a variety of sulfur-containing “chemotypes” and products with different formation pathways, functionalities, and physical and chemical reactivities. The most biologically relevant cysteine-modified intermediates and products found *in vivo* are shown in **Figure 1.1**.



**Figure 1.1.** Examples of biologically relevant cysteine functional modification intermediates and products.<sup>7</sup>

In recent years, S-sulfhydration (-S—SH) has gained considerable attention among cysteine-derived post-translational modifications. Due to its proposed role in redox signaling, cysteine sulfhydration has been suggested as an important modification to form

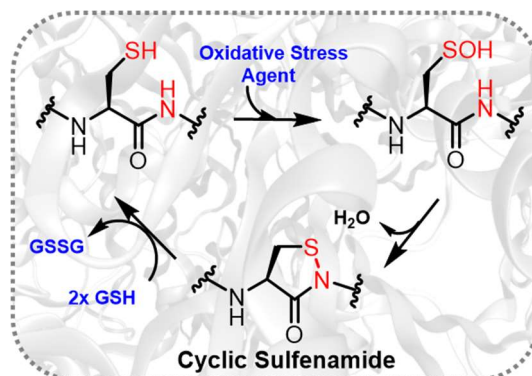
the persulfide bond (Cys-S—SH).<sup>15-16</sup> As various studies have shown, S-sulfhydration can result from H<sub>2</sub>S-mediated cysteine modification.<sup>15-18</sup> Other oxidative cysteine post-translational modifications (for instance formation of disulfides, sulfenic acids, or nitrosothiols) can also undergo further reactions with H<sub>2</sub>S to form cysteine persulfide. Furthermore, recent studies have shown that other reactive sulfur species (RSS) in the form of hydrogen persulfide (H<sub>2</sub>S<sub>2</sub>) or polysulfides (R—S<sub>n</sub>H,  $n > 2$ ) can also yield cysteine persulfide.<sup>19</sup>

Cysteine persulfide has some distinct yet important chemical properties compared to the thiol form, such as different S—H bond energies and greater stability of the perthiyl radical (RSS<sup>•</sup>) *versus* the thiyl radical (RS<sup>•</sup>).<sup>20</sup> Notably, cysteine persulfide with a pK<sub>a</sub> of ~6.5 is a stronger acid relative to cysteine with a pK<sub>a</sub> of ~8.5.<sup>21-22</sup> Thus, Cys-SS<sup>-</sup> exists in a deprotonated state and behaves as a better nucleophile at physiological pH. However, much of the biological chemistry of Cys-SSH remains unknown due to its higher reactivity and inherent instability.

In addition to S-sulfhydration, covalent sulfur-nitrogen bonds (S—N) are also found in many post-translationally modified cysteine residues (see **Figure 1.2**). Sulfenylamides (sulfenamides) contain an S—N bond between divalent sulfur and trivalent nitrogen atoms. The most notable instance of a naturally occurring sulfenamide is the cyclic form of thiazolidine-3-one obtained from the reaction of cysteine sulfenic acid and the nitrogen atom of a neighboring amino acid.<sup>23-25</sup> Interestingly, formation of a cyclic sulfenylamide is considered a cellular protective mechanism against oxidative stress such as the irreversible formation of cysteine sulfinic (Cys-SO<sub>2</sub>H) and sulfonic acid (Cys-SO<sub>3</sub>H), both of which lead to protein degradation and cell death.<sup>4, 23, 26-27</sup> As shown in **Figure 1.2**, reduction of cyclic sulfenylamide through further reaction with biological thiols such as glutathione (GSH) reverses the sulfenylamide to the catalytically active form of cysteine.

Despite the high specificity, functionality and the biological importance of compounds that contain S—N bonds, they are less abundant in biological systems. This scarcity of

S—N bonds in biological contexts prompts further study into their occurrence, formation, and chemistry.



**Figure 1.2.** The protective role of S—N bond formation due to oxidative stress in biological systems.<sup>28</sup>

### 1.3 Catalytic Solvents and Proteins

Not surprisingly, thousands of (bio)chemical processes take place in the solution phase or a liquid environment. The solution phase is defined as a homogenous medium in which (bio)chemical reactants encounter, impact, and react with each other through several short- and long-range interactions. Pioneering research by Menshutkin on the quaternization of triethylamine with iodoethane demonstrated how the reaction rate depends significantly on the choice of solvent.<sup>29</sup> Today, studies have proven that not only the reaction rate, but also many physical and chemical molecular properties, structures, stability, mechanisms, the overall kinetics and thermodynamics of the reaction, and most importantly, the consequences of the reaction can be remarkably influenced in the solution phase as a result of solvent-solute interactions.<sup>30</sup>

In addition to many organic solvents with different polarities and ability to act as either a hydrogen bond donor or acceptor, proteins are also classified as a powerful environment for catalyzing many amazing biomolecular processes. One of the tremendous features of a protein environment is the facilitation of biochemical reactions. Enzymes, proteins with catalytic properties, are combinations of sequenced amino acids, (in)organic cofactors,



metals, and thousands of water molecules. They are capable of accelerating reaction rates by lowering the activation barrier.<sup>31</sup> During an enzymatic process, reaction rates are dramatically increased, occurring in fractions of seconds instead of years under the compatible conditions of life (temperature and pressure) without losing their catalytic properties, being consumed in the reaction, or changing reaction equilibrium constants. This results from highly reactive transition state stabilization, substrate destabilization, conformational changes, electrostatic and van der Waals interactions, and expanded H-bonding networks in the enzyme active sites.<sup>32-36</sup> Alternatively, they can alter the reaction pathway by catalyzing a lower-energy direction. Therefore, the catalytic behavior of proteins makes them a crucial and inseparable part of living organisms in controlling and performing processes required for maintaining life.

### **1.4 Computational Chemistry's Application**

To date, computational chemistry has been established as an invaluable tool to investigate several (bio)chemical systems. It is widely used to study, model and simulate chemical properties and behavior, functionality, reactivity, and mechanisms of several reactions in varied-sized systems; from a single atom to thousands of atoms in biomolecular contexts.<sup>37-39</sup> Computational chemistry has been applied to reveal unknown and novel chemistry in recent years, in combination with experiment or even beyond it, where experimental chemistry fails to find an answer. A broad range of computational modeling techniques has been developed over time, each with its own strengths and shortcomings and each suitable for different types of systems.<sup>40</sup> For instance, modeling a single bond formation in a diatomic system, exploring kinetic and thermodynamic properties in multi-atom systems, and elucidating complex enzymatic processes in today's computational world is feasible and accessible applying the proper computational techniques.<sup>37-39, 41-42</sup>

It should be stated that choosing the best-suited computational method and tool to the system of interest to provide reliable and accurate results is still a challenging task for many

computational chemists despite substantial progress over the past few decades.<sup>43-44</sup> Therefore, assessing the suitability and compatibility of computational methods is essential to achieve meaningful conclusions with high accuracy and reliability. Performing systematic benchmark studies is a way to overcome this uncertainty and validate the applied methods which then can efficiently be expanded to the larger models of interest.<sup>45-49</sup> Furthermore, referring to the literature is highly recommended to make an informed choice of computational approaches.

In this dissertation, following an assessment study to find the most accurate structural and thermochemical data on a number of bio-relevant sulfur-containing compounds, a multi-scale computational approach was applied to explore some sulfur-containing systems, mostly involving sulfur-nitrogen bond species and persulfide derivatives.

### 1.5 References

1. Block, E., The Organosulfur Chemistry of the Genus *Allium* – Implications for the Organic Chemistry of Sulfur. *Angew. Chem. Int. Ed.* **1992**, *31*, 1135-1178.
2. Kim, H.-J.; Ha, S.; Lee, H. Y.; Lee, K.-J., ROSics: Chemistry and Proteomics of Cysteine Modifications in Redox Biology. *Mass Spectrom. Rev.* **2015**, *34*, 184-208.
3. Reddie, K. G.; Carroll, K. S., Expanding the Functional Diversity of Proteins Through Cysteine Oxidation. *Curr. Opin. Chem. Biol.* **2008**, *12*, 746-754.
4. Gupta, V.; Carroll, K. S., Sulfenic acid Chemistry, Detection and Cellular Lifetime. *BBA-Gen. Subjects* **2014**, *1840*, 847-875.
5. Petkowski, J. J.; Bains, W.; Seager, S., Natural Products Containing a Nitrogen-Sulfur Bond. *J. Nat. Prod.* **2018**, *81*, 423-446.
6. Waldman, A. J.; Ng, T. L.; Wang, P.; Balskus, E. P., Heteroatom-Heteroatom Bond Formation in Natural Product Biosynthesis. *Chem. Rev.* **2017**, *117*, 5784-5863.
7. Go, Y.-M.; Chandler, J. D.; Jones, D. P., The Cysteine Proteome. *Free Radic. Bio. Med.* **2015**, *84*, 227-245.

8. Paulsen, C. E.; Carroll, K. S., Cysteine-Mediated Redox Signaling: Chemistry, Biology, and Tools for Discovery. *Chem. Rev.* **2013**, *113*, 4633-4679.
9. Couturier, J.; Chibani, K.; Jacquot, J.-P.; Rouhier, N., Cysteine-Based Redox Regulation and Signaling in Plants. *Front. Plant Sci.* **2013**, *4*, 105.
10. Cuello, F.; Eaton, P., Cysteine-Based Redox Sensing and Its Role in Signaling by Cyclic Nucleotide-Dependent Kinases in the Cardiovascular System. *Annu. Rev. Physiol.* **2019**, *81*, 63-87.
11. Chung, H. S.; Wang, S.-B.; Venkatraman, V.; Murray, C. I.; Van Eyk, J. E., Cysteine Oxidative Posttranslational Modifications: Emerging Regulation in the Cardiovascular System. *Circ. Res.* **2013**, *112*, 382-392.
12. Harmel, R.; Fiedler, D., Features and Regulation of Non-Enzymatic Post-Translational Modifications. *Nat. Chem. Biol.* **2018**, *14*, 244-252.
13. Walsh, G., Post-Translational Modifications of Protein Biopharmaceuticals. *Drug Discov. Today* **2010**, *15*, 773-780.
14. Carrico, I. S., Chemoselective Modification of Proteins: Hitting the Target. *Chem. Soc. Rev.* **2008**, *37*, 1423-1431.
15. Ida, T.; Sawa, T.; Ihara, H.; Tsuchiya, Y.; Watanabe, Y.; Kumagai, Y.; Suematsu, M.; Motohashi, H.; Fujii, S.; Matsunaga, T.; Yamamoto, M.; Ono, K.; Devarie-Baez, N. O.; Xian, M.; Fukuto, J. M.; Akaike, T., Reactive Cysteine Persulfides and S-Polythiolation Regulate Oxidative Stress and Redox Signaling. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 7606-7611.
16. Yadav, P. K.; Martinov, M.; Vitvitsky, V.; Seravalli, J.; Wedmann, R.; Filipovic, M. R.; Banerjee, R., Biosynthesis and Reactivity of Cysteine Persulfides in Signaling. *J. Am. Chem. Soc.* **2016**, *138*, 289-299.
17. Millikin, R.; Bianco, C. L.; White, C.; Saund, S. S.; Henriquez, S.; Sosa, V.; Akaike, T.; Kumagai, Y.; Soeda, S.; Toscano, J. P.; Lin, J.; Fukuto, J. M., The Chemical

Biology of Protein Hydropersulfides: Studies of a Possible Protective Function of Biological Hydropersulfide Generation. *Free Radic. Bio. Med.* **2016**, *97*, 136-147.

18. Mishanina, A. V.; Libiad, M.; Banerjee, R., Biogenesis of Reactive Sulfur Species for Signaling by Hydrogen Sulfide Oxidation Pathways. *Nat. Chem. Biol.* **2015**, *11*, 457-464.

19. Iciek, M.; Kowalczyk-Pachel, D.; Bilska-Wilkosz, A.; Kwiecien, I.; Gorny, M.; Wlodek, L., S-Sulfhydration as a Cellular Redox Regulation. *Biosci. Rep.* **2016**, *36*, 1-17.

20. Ono, K.; Akaike, T.; Sawa, T.; Kumagai, Y.; Wink, D. A.; Tantillo, D. J.; Hobbs, A. J.; Nagy, P.; Xian, M.; Lin, J.; Fukuto, J. M., Redox Chemistry and Chemical Biology of H<sub>2</sub>S, Hydropersulfides, and Derived Species: Implications of Their Possible Biological Activity and Utility. *Free Radic. Bio. Med.* **2014**, *77*, 82-94.

21. Park, C. M.; Weerasinghe, L.; Day, J. J.; Fukuto, J. M.; Xian, M., Persulfides: Current Knowledge and Challenges in Chemistry and Chemical Biology. *Mol. Biosys.* **2015**, *11*, 1775-1785.

22. Lau, N.; Pluth, M. D., Reactive Sulfur Species (RSS): Persulfides, Polysulfides, Potential, and Problems. *Curr. Opin. Chem. Biol.* **2019**, *49*, 1-8.

23. Sarma, B. K.; Mugesh, G., Redox Regulation of Protein Tyrosine Phosphatase 1B (PTP1B): A Biomimetic Study on the Unexpected Formation of a Sulfenyl Amide Intermediate. *J. Am. Chem. Soc.* **2007**, *129*, 8872-8881.

24. Defelipe, L. A.; Lanzarotti, E.; Gauto, D.; Marti, M. A.; Turjanski, A. G., Protein Topology Determines Cysteine Oxidation Fate: The Case of Sulfenyl Amide Formation among Protein Families. *Plos Comput. Biol.* **2015**, *11*, 1-25.

25. Forman, H. J.; Davies, M. J.; Krämer, A. C.; Miotto, G.; Zaccarin, M.; Zhang, H.; Ursini, F., Protein Cysteine Oxidation in Redox Signaling: Caveats on Sulfenic Acid Detection and Quantification. *Arch. Biochem. Biophys.* **2017**, *617*, 26-37.

26. Roos, G.; Messens, J., Protein Sulfenic Acid Formation: From Cellular Damage to Redox Regulation. *Free Radic. Bio. Med.* **2011**, *51*, 314-326.

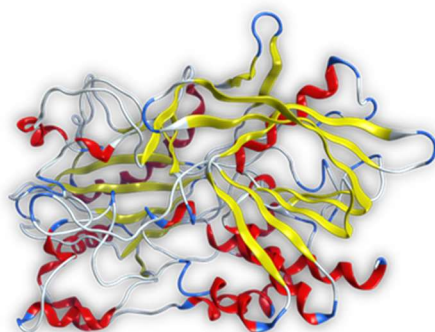
27. Groitl, B.; Jakob, U., Thiol-Based Redox Switches. *BBA-Proteins Proteom.* **2014**, *1844*, 1335-1343.
28. Petkowski, J. J.; Bains, W.; Seager, S., Natural Products Containing a Nitrogen–Sulfur Bond. *J. Nat. Prod.* **2018**, *81*, 423-446.
29. Reinheimer, J. D.; Harley, J. D.; Meyers, W. W., Solvent Effects in the Menschutkin Reaction. *J. Org. Chem.* **1963**, *28*, 1575-1579.
30. Reichardt, C., *Solvents and Solvent Effects in Organic Chemistry*. Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, 2010.
31. Wolfenden, R.; Snider, M. J., The Depth of Chemical Time and the Power of Enzymes as Catalysts. *Acc. Chem. Res.* **2001**, *34*, 938-945.
32. Giraldo, J.; Roche, D.; Rovira, X.; Serra, J., The Catalytic Power of Enzymes: Conformational Selection or Transition State Stabilization? *FEBS Lett.* **2006**, *580*, 2170-2177.
33. Zhang, X.; Houk, K. N., Why Enzymes Are Proficient Catalysts: Beyond the Pauling Paradigm. *Acc. Chem. Res.* **2005**, *38*, 379-385.
34. Richard, J. P., Protein Flexibility and Stiffness Enable Efficient Enzymatic Catalysis. *J. Am. Chem. Soc.* **2019**, *141*, 3320-3331.
35. Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M., Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **2006**, *106*, 3210-3235.
36. Benkovic, S. J.; Hammes-Schiffer, S., A Perspective on Enzyme Catalysis. *Science* **2003**, *301*, 1196-1202.
37. Gherib, R.; Dokainish, H. M.; Gault, J. W., Multi-Scale Computational Enzymology: Enhancing Our Understanding of Enzymatic Catalysis. *Int. J. Mol. Sci.* **2014**, *15*, 401-422.
38. Cheng, G.-J.; Zhang, X.; Chung, L. W.; Xu, L.; Wu, Y.-D., Computational Organic Chemistry: Bridging Theory and Experiment in Establishing the Mechanisms of Chemical Reactions. *J. Am. Chem. Soc.* **2015**, *137*, 1706-1725.

39. Jinich, A.; Flamholz, A.; Ren, H.; Kim, S.-J.; Sanchez-Lengeling, B.; Cotton, C. A. R.; Noor, E.; Aspuru-Guzik, A.; Bar-Even, A., Quantum Chemistry Reveals Thermodynamic Principles of Redox Biochemistry. *Plos Comput. Biol.* **2018**, *14*, 1-22.
40. Martin-Santamaria, S., *Computational Tools for Chemical Biology*. Royal Society of Chemistry: London, UK: 2018.
41. Hidalgo, A.; Giroday, T.; Mora-Diez, N., Thermodynamic Stability of Neutral and Anionic PFOAs. *Theor. Chem. Acc.* **2015**, *134*, 124.
42. Ramis, R.; Casasnovas, R.; Ortega-Castro, J.; Frau, J.; Álvarez-Idaboy, J. R.; Mora-Diez, N., Modelling the Repair of Carbon-Centred Protein Radicals by the Antioxidants Glutathione and Trolox. *New J. Chem.* **2019**, *43*, 2085-2097.
43. Carvalho, A. T. P.; Barrozo, A.; Doron, D.; Kilshtain, A. V.; Major, D. T.; Kamerlin, S. C. L., Challenges in Computational Studies of Enzyme Structure, Function and Dynamics. *J. Mol. Graph. Model.* **2014**, *54*, 62-79.
44. Bachrach, S. M., Challenges in Computational Organic Chemistry. *WIREs. Comput. Mol. Sci.* **2014**, *4*, 482-487.
45. Simm, G. N.; Proppe, J.; Reiher, M., Error Assessment of Computational Models in Chemistry. *Chimia* **2017**, *71*, 202-208.
46. Mata, R. A.; Suhm, M. A., Benchmarking Quantum Chemical Methods: Are We Heading in the Right Direction? *Angew. Chem. Int. Ed.* **2017**, *56*, 11011-11018.
47. Nikoo, S.; Meister, P. J.; Hayward, J. J.; Gault, J. W., An Assessment of Computational Methods for Calculating Accurate Structures and Energies of Bio-Relevant Polysulfur/Selenium-Containing Compounds. *Molecules* **2018**, *23*, 3323.
48. Baciú, C.; Gault, J. W., An Assessment of Theoretical Methods for the Calculation of Accurate Structures and S–N Bond Dissociation Energies of S-Nitrosothiols (RSNOs). *J. Phys. Chem. A* **2003**, *107*, 9946-9952.

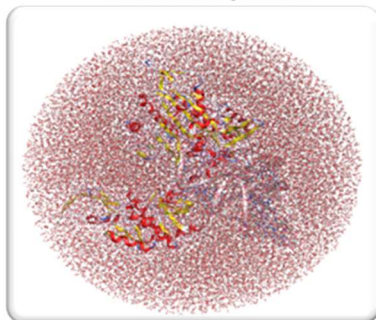
49. Brás, N. F.; Perez, M. A. S.; Fernandes, P. A.; Silva, P. J.; Ramos, M. J., Accuracy of Density Functionals in the Prediction of Electronic Proton Affinities of Amino Acid Side Chains. *J. Chem. Theory Comput.* **2011**, 7, 3898-3908.

## CHAPTER 2

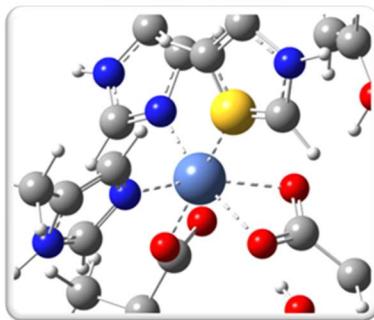
# Computational Methods



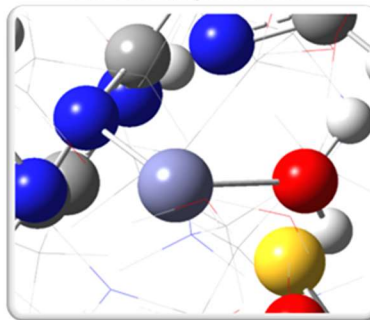
**Molecular Dynamics**



**QM-Cluster Model**



**ONIOM: QM/MM Model**





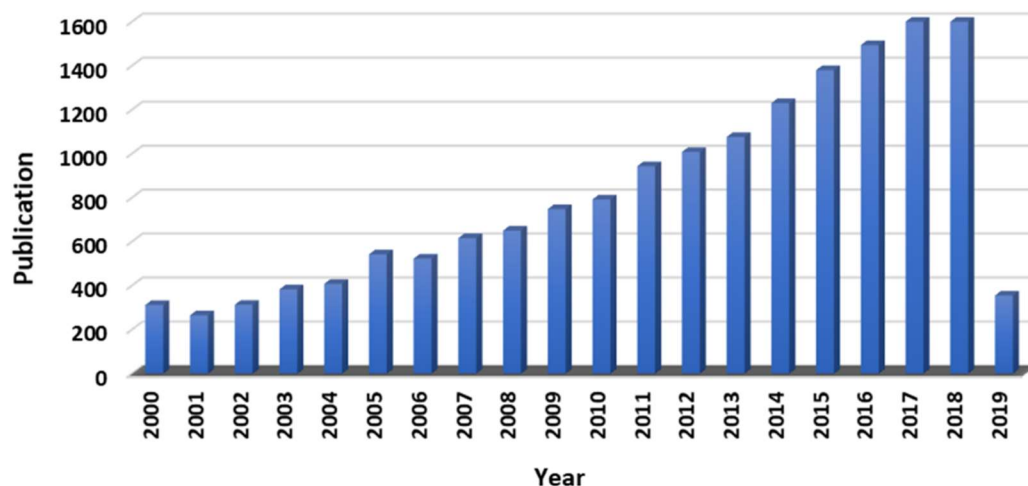
### 2.1 Computational Chemistry

Beginning with Planck's quantum theory in the 1950s, computational chemistry and biochemistry have improved impressively; from a rarity to a powerful partner with experimental (bio)chemistry.<sup>1</sup> In fact, computational chemistry is a general term, covering any use of computation to investigate (bio)chemical systems. Today, the chief interests of the (bio)chemist are to design structures and reactions, elucidate the functionality, reactivity, mechanism, and importantly, selecting chemical properties through a series of rational modifications. The one potent tool which does not involve the study of the chemical context in real time and physical matter as such is computational chemistry. It uses computers to apply the mathematical equations and theorems of quantum chemistry to study (bio)chemical systems, model and simulate molecular behavior and properties, and find solutions for challenging experimental problems. Owing to the development of the computational software and hardware as well as theoretical chemistry over the past few decades, computational chemistry has been expanded to a broad field of chemistry. It ranges extensively from (in)organic synthesis, drug design, and pharmaceuticals to the study the mechanistic function of macromolecules such as enzymes. **Figure 2.1** indicates the growth rate of computational chemistry in the 21<sup>st</sup> century.<sup>2</sup>

Over time, numerous computational methods and approaches have been developed. Indeed, a variety of methods are now available for systems of differing complexities and sizes ranging from a single or a few atoms to thousands of atoms. In general, these methods include density functional theory (DFT), molecular mechanics and dynamics (MM and MD), quantum mechanics (QM)-cluster and quantum mechanics/molecular mechanics (QM/MM). Individual or combination of the methods alongside experimental data enables deeper chemical insight into the target.

In this dissertation, multi-scale computational methods have been applied to study some (bio)molecular systems. Many excellent books and publications have discussed these

methods and approaches.<sup>3-7</sup> Hence, only some of the key features of applied methods and central relevant theories in this dissertation are briefly discussed herein.

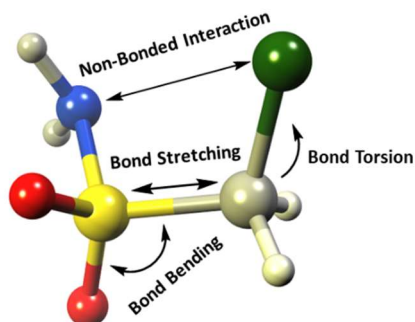


**Figure 2.1.** Number of publications per year; a result of the search on ISI Web of Science using the topic of “Computational Chemistry” (accessed on 28<sup>th</sup> March 2019).

## 2.2 Molecular Mechanics

Classical mechanics is one of the common approaches to study molecular constructions. The resulting type of method is referred to as molecular mechanics (MM), also known as force field methods.<sup>3-5</sup> Indeed, molecular mechanics as an empirical-based approach is categorized among the simplest and most inexpensive computational methodologies in today's computational world. Electronic motions are ignored in mechanics calculation thus, the total energy of the system is a function of solely nuclear coordinates, made possible by the Born-Oppenheimer approximation.<sup>3</sup> Consequently, this method cannot be applied to describe bond forming and breaking processes or calculating electronic properties such as UV-Vis spectra. In other words, MM methods are only able to describe interactions between nuclei, both covalent bonded and non-bonded interactions. In the case of covalently bonded interaction, the energy of the system is a description of bond stretching,

bond bending, and torsion angle (**Figure 2.2**). Also, the non-bonded interactions are described by both van der Waals and electrostatic interactions.



**Figure 2.2.** Representation of the bonded and non-bonded interactions calculated by molecular mechanics.

Thus the total energy of the MM system can be mathematically represented as in **Equation 2.1**. Where the first three terms describe the energy components of bonded interactions: bond stretching, bending, and torsion, respectively. The last two terms are non-bonded van der Waals and electrostatic contributions which are handled using Lennard-Jones potentials and Colomb's law, respectively.

$$U_{Total} = U_{Str} + U_{Bend} + U_{Tors} + U_{vdW} + U_{Elec} \quad (\text{Equation 2.1})$$

As mentioned previously, molecular mechanics is an empirical method and uses constant parameters obtained from experimental data, for instance, crystal structures or high accuracy *ab initio* calculations. A set of mathematical expressions and constant parameters form a particular force field (FF). Many empirical force fields have been developed, each specifically parametrized for particular use in different kinds of systems from organic to biomolecular contexts. Developing a proper and accurate force field is still a challenging and demanding task, especially for those biochemical systems including transition metal ions (*e.g.*, Mn and Mo) with multiple oxidation states and other inorganic cofactors.<sup>8-11</sup> Thus, to set up molecular mechanics, it is important to know the system of

interest and correctly select the best-suited force field. Assisted Model Building with Energy Refinement (AMBER)<sup>12</sup>, Chemistry at Harvard Macromolecular Mechanics (CHARMM)<sup>13</sup>, and Groningen molecular simulation GROMOS<sup>14</sup> are the most popular and well-adapted force fields in modeling biomolecular macromolecules. It is worth stating that observing slight differences between molecular mechanics calculations applying different force fields for the same system is not surprising since each atom has its own defined parameters and definitions that vary across the force fields.

### 2.3 Molecular Dynamics (MD) Simulations

Molecular dynamics is a simulation of the time-dependent behavior of molecular systems at atomic resolution.<sup>3-5</sup> In fact, molecular dynamics generates serial configurations of changes in a system with respect to time. Thus, it can be applied successfully as a fundamental and effective tool for modeling and elucidating dynamics of many (bio)chemical pathways. It gives deep insight into molecular behaviors and changes over a fixed period of time and is particularly useful for biomolecular systems containing tens of thousands of atoms. Indeed, MD simulations have been established as a vital technique in computational enzymology to investigate behaviors including conformational interconversions and protein interactions with ligands, inhibitors, and substrates. Most commonly, Newton's second law of motion is used to predict molecular trajectories in which the required forces and potential energies are computed *via* molecular mechanics force fields. In fact, similar to MM, molecular dynamics is an empirical-based approach where it uses parametrized particular force fields for simulations. To date, many molecular modeling software packages have been developed including NAMD<sup>15</sup>, GROMACS<sup>16</sup>, AMBER<sup>17</sup>, GROMOS<sup>18</sup>, and CHARMM<sup>19</sup> result of modern simulation algorithms and advanced computational software and hardware. Many of the common force fields are available within these MD software packages. In this dissertation, molecular dynamics simulations have been performed using the NAMD program and the AMBER force field.

As mentioned the result of molecular dynamics simulation is molecular trajectory, a series of structural changes as a function of time. Analyzing the generated trajectories can yield the desired information about the system of interest such when it reaches equilibrium, the snapshots should be selected as the initial templates for further QM-cluster and QM/MM calculations.

## 2.4 Quantum Mechanical (QM) Methods

To date, the two major quantum mechanical methods used to calculate the electronic structures of thousands of molecular systems include *ab initio* (wavefunction-based) and density functional theory (DFT). These methods endeavor to provide an approximate solution to solve the Schrödinger equation, the heart of quantum chemistry. The field-free Hamiltonian of the time-independent, non-relativistic Schrödinger equation of a molecular system is fully shown in **Equation 2.2**.<sup>3, 20</sup>

$$\hat{H} = - \sum_{i=1}^{N_E} \frac{1}{2} \nabla_i^2 - \sum_{A=1}^{N_N} \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^{N_E} \sum_{A=1}^{N_N} \frac{Z_A}{r_{iA}} + \sum_{A=1}^{N_N} \sum_{B>A}^{N_N} \frac{Z_{AB}}{r_{AB}} + \sum_{i=1}^{N_E} \sum_{j>i}^{N_E} \frac{1}{r_{ij}} \quad (\text{Equation 2.2})$$

The first two terms are the kinetic energy of  $N_E$  electrons and  $N_N$  nuclei, respectively. The following three terms correspond to the potential energy of electron-nuclei attraction, nuclei-nuclei, and electron-electron repulsions, respectively.  $M_A$  and  $Z_A$  are the mass and the atomic number of nuclei, respectively. The distance between the two particles is shown by  $r$ .

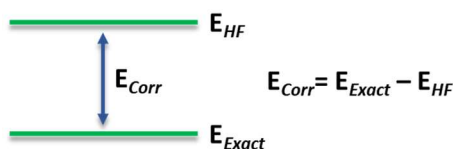
To simplify the complexity of the above Hamiltonian for multi-electron systems some approximations can be applied. The Born-Oppenheimer (BO) approximation and orbital approximation are the two most common ones. Due to heavier masses and much slower movement of nuclei relative to the electrons ( $\sim 1800$  times more massive than electrons), nuclei are considered as stationary particles. Therefore, electrons are assumed to be moving in a field of fixed nuclei (BO approximation). Further simplification can also be made

through the orbital approximation, which states electrons move independently of each other. In other words, in a multi-electron system, each individual electron has its own one-electron orbital. More simply, the orbital approximation ignores the Coulombic repulsive interactions between electrons, despite their interactions in a real system!<sup>3, 20</sup>

### 2.4.1 *Ab Initio* Based Methods

Hartree-Fock (HF), the simplest *ab initio* method, is established based on moving each electron in a smeared field generated by the rest of the electrons. By ignoring explicit repulsive interactions, this method only takes into account the average effect of Coulombic electron-electron repulsion. Excluding the electron correlation (repulsion) energy in the HF method leads to overestimating the total electronic energy of the system and significantly affect molecular and electronic properties such as shortening the covalent bonds and/or lengthening the electrostatic interactions since there are not enough contributions between electrons.

**Figure 2.3** represents the difference between the limiting HF energy and the exact non-relativistic energy called “correlation energy”. Obviously, the correlation energy is always negative with a stabilizing effect.



**Figure 2.3.** Schematic representation of correlation energy calculation.

To account for electron correlation into the computed electronic energy and geometrical parameters to overcome the HF limitations, several methods have also been developed including correlated *ab initio* methods.<sup>3</sup> One of these methods uses configuration interactions (CI method). In this approach, the HF wavefunction is improved by the addition of a determined number of excited configurations. The CI calculation with all the

possible excitations is called full configuration interaction (FCI) which is the exact solution to the Schrödinger equation. The way to limit the FCI and make a practical approach is truncation. For instance, CID and CISD are limited to double only and single and double excitations, respectively. Although they suffer from size-consistency which makes errors unpredictable with larger system size. To overcome the lack of the size-consistency of CISD some improved methods were developed such as quadratic configuration interaction (QCI). It modifies the size-consistency error by addition of quadratic terms to a linear extension of the CID and CISD wavefunction. For instance, QCISD is a contribution of all singly- and doubly-excited configurations with some quadruply-excited configurations.

Despite accurate and reliable results of CI methods compared to HF, these methods are extremely computationally expensive and thus unfeasible for large systems such as enzymes. Therefore, they can be employed as a benchmark calculation against other quantum mechanical methods especially QCISD in small-sized systems (**Chapter 3**).<sup>21-22</sup>

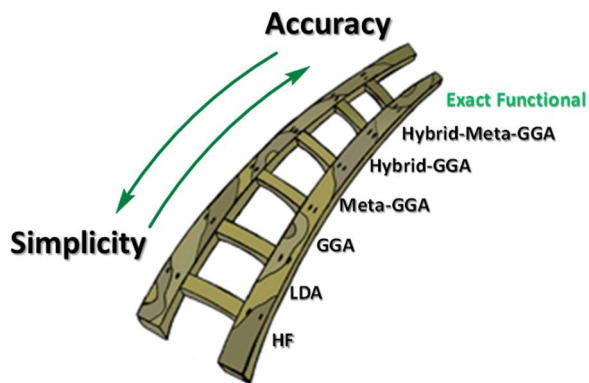
### 2.4.2 Density Functional Theory (DFT)-Based Methods

To overcome the limitations of the wavefunction-based methods, the tremendous effort of Kohn and Pople was awarded with the Nobel Prize in chemistry in 1998 for "developing density functional theory (DFT)".<sup>23</sup> Indeed, density functional theory with the aim of application to large systems, accounts for the electron correlation energy and reduces the computational costs. Surprisingly, the results obtained from DFT-based methods can compete with those of correlated *ab initio* methods in both accuracy and specifically computational cost.

The main story behind the success of DFT is that the energy of a multi-electron system is a function of the electron density ( $\rho$ ) instead of the wavefunction.<sup>3, 20, 24</sup> Indeed, in DFT calculations, molecular and electronic properties of a multi-electron system can be determined by electron density. The density function only relies upon three spatial coordinates ( $x$ ,  $y$ , and  $z$ ) for each electron ( $3N$ ), regardless of the number of electrons. In

contrast, wavefunction-based methods depend on  $3N$  spatial and  $N$  spin coordinates ( $4N$  in total). All these dramatically result in a reduction of complexity and computational costs and speeding up calculations. For such reasons, DFT functionals are able to be applied efficiently in studying large-sized chemical systems.

Notably, the exact form of the exchange-correlation (XC) energy for density functionals is not known and several approximate functionals have been developed over the time. Today, a large number of DFT functionals exist which differs in expressing exchange-correlation functionals. Some only apply parameters based on the fundamental physical constants while others use empirical parameters fitted to experimental data and high accuracy QM methods. **Figure 2.4** outlines the Perdew's '*Jacob ladder*' of approximations for the exchange-correlation energy, which has been classified based on the different level of XC approximations as a function of electron density.<sup>25</sup> It begins with the Hartree world with no exchange-correlation toward the chemical accuracy.



**Figure 2.4.** Schematic representation of the Perdew's '*Jacob ladder*' for 5-generated DFT functionals.<sup>22</sup>

After the no exchange-correlation Hartree world, the first rung of the Jacob ladder starts with the local density approximation (LDA) where the energy is locally dependent only upon the electron density.<sup>26</sup> It is the simplest, but the least impressive approach to solving the exchange-correlation problem since it has a tendency to exaggerate the bond strength



and shorten the corresponding bond length. In contrast, generalized gradient approximation (GGA) methods remarkably improve over the LDA methods by including the first derivative of electron density (gradient of density) and gives more accurate results. There are several different parameterizations for the GGA, some from empirical parameters or from first principles. The most popular gradient-corrected exchange and correlation functionals at the present time include Becke's 1988 functional (B88)<sup>27</sup>, Perdew and Wang's 1991 functional (PW91)<sup>28</sup> and Lee-Yang-Parr functional (LYP)<sup>29</sup>. As well, *meta*-generalized gradient approximation (*meta*-GGA) functionals go beyond the first-order density by the inclusion of the kinetic energy density.<sup>30</sup> Indeed, in *meta*-GGA methods the exchange-correlation energy relies upon the second derivative of the electron density.

Moving to higher rungs of the Jacob ladder, Hybrid-GGA functionals calculate a portion of the exchange-correlation energy using the HF theory.<sup>31</sup> The result of this combination method to obtain so-called exact exchange-correlation energy is known as hybrid DFT methods.

Among many developed hybrid functionals in recent years, the B3LYP exchange-correlation functional is one of the most successful and widely used and is significantly responsible for DFT becoming one of the trendiest tools in the computational chemistry community.<sup>32</sup> Its surprisingly good performance in many chemical applications was a turning point for DFT to be accepted as the most common electronic structure method in computational chemistry. As shown in **Equation 2.3**, the B3LYP exchange-correlation energy expression is a combination of LDA, GGA functionals and HF exchange parameters.<sup>29, 31</sup>

$$E_{XC}^{B3LYP} = (1 - a)E_X^{LDA} + aE_X^{HF} + b\Delta E_X^B + (1 - c)E_C^{LDA} + cE_C^{LYP} \quad (\text{Equation 2.3})$$

Where  $a=0.1161$ ,  $b=0.9262$ , and  $c=0.8133$ .

Although it suffers from some performance challenges.<sup>3, 33-34</sup> These issues briefly stem from self-interaction error (electron interaction with itself), ignoring non-covalent long-range interaction, and the inability of hybrid DFT in describing systems containing transition metals.

The newly developed Minnesota density functionals, M06 and its improved version M08 classes of functionals, are mostly based on the Hybrid-*meta*-GGA or *meta*-GGA which have been shown to overcome the aforementioned difficulties.<sup>35-37</sup> These highly parametrized functionals with kinetic energy dependence also differ in the percentage of HF exchange, which makes them the most broadly adopted of recently developed functionals in recent years.<sup>38</sup> For instance, M06-2X as a hybrid-*meta*-functional with 54% HF exchange performs well for studying main group thermochemistry, non-covalent, long-range interactions, and kinetics while is not recommended for systems including transition metal ions. On the other hand, the local M06-L functional, with no HF exchange contribution is the most accurate one for transition metals, inorganic and metalorganic chemistry.

As observed, moving to the higher rungs of Jacob ladder, both accuracy and complexity of exchange-correlation components increase considerably to (hopefully) achieve to the exact energy. Although one should keep in mind that DFT methods were developed more recently than *ab initio* methods. Thus, there are some critical issues that need careful considerations using DFT methods such as weakness in London dispersion forces description and incapacity for systematic functional improvement.

Despite the diversity in available functionals, choosing the best-adapted quantum mechanics method is not an easy task. One should consider all chemical aspects and the nature of the system of interest to decide the QM-functional. Referring to previous literature and conducting a systematic benchmarking study provide an opportunity to make the best choice.

## 2.5 Basis Set

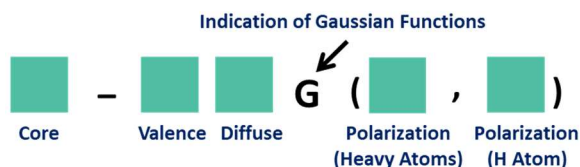
Aside from choosing the QM method, choosing a proper basis set has a critical role in achieving robust accuracy. The basis set is defined as the set of the mathematical functions to represent the shape of an atomic orbital.<sup>3, 20</sup> Indeed, the basis set is a linear combination of basis functions to describe the electronic wavefunction in the QM methods. Today, Gaussian-type basis functions (GTFs) are by far the most commonly used basis sets due to faster and simpler computational calculations. Expectedly, better accuracy in the orbital description is attained by applying more basis functions. There are several different ways to extend basis sets by including a combination of split-valence, polarized, and diffuse functions.

***Split-Valence:*** Since the valence electrons are involved in most molecular bonding, two or more basis functions can be applied to accurately describe the valence atomic orbitals. As a result, valence atomic orbitals in molecules can expand and adjust for variable atom size. Double-Zeta (DZ), *e.g.*, 6-31G and Triple-Zeta (TZ), *e.g.*, 6-311G include two and three basis functions for each valence orbital description, respectively. It is worth mentioning that 6 represents the number of Gaussian functions to describe the inner shell orbitals (core orbitals).

***Polarized Functions:*** Such functions allow orbitals to be polarized with the higher orbital angular momentum. Thus, it detracts orbitals from their original atomic symmetry and adds flexibility within basis sets. *p*- and *d*- polarization functions can be applied to the hydrogen atom, while *d*- and *f*- are common for heavy atoms (all atoms but H atom).

***Diffuse Function:*** To gain more accurate descriptions of atomic orbitals, especially for systems including anions, weak bonds, and long-range interactions, employing diffuse functions are important. The (+) sign is a representation of the diffuse functions only on heavy atoms, while (++) applies diffuse functions on both heavy and non-heavy atoms.

The Pople basis set family is the most widely used with high accuracy in the obtained results. The composition of a Pople basis set is shown in **Figure 2.5**.



**Figure 2.5.** Schematic representation of Pople basis sets compositions.

For instance, the 6-311+G(d,p) basis set is an example of a split-valence triple-zeta basis set which includes diffuse functions on only heavy atoms, *d*-polarized functions on heavy atoms, and *p*-polarized functions on hydrogens atoms. Pople basis sets are the only basis set used in this dissertation.

Like QM methods, choosing the right basis set requires careful consideration of the desired system. Performing a basis set assessment (see **Chapter 3**), referring to the literature, and finding a similar system to one's interest are recommended before choosing the basis set. It is worth mentioning both choosing a basis set and QM method are extremely important in determining the accuracy of the chemical models and eventually achieving more accurate and reliable results.

## 2.6 Quantum Mechanical (QM)-Cluster Approach

Quantum Mechanical (QM) calculations have been identified as one of the straightforward and successful methods for mechanistic investigations in a broad spectrum of (bio)chemical systems, particularly enzymatic studies. In recent years many different systems have been explored using these methods and have resulted in valuable insights.<sup>39-</sup>

<sup>44</sup> The QM-cluster only approach mainly focuses on the small, but key area of the reaction center or small area surrounding the enzyme active site. Indeed, the small size of the selected model permits applying highly accurate QM methods to study the chemical reaction. This methodology depends significantly on the system size due to increased

computational cost. Large systems cannot be treated quantum mechanically with the most accurate computational approaches.

As a matter of fact, the selection of the QM region has a substantial impact on the obtained findings and results. Some basic criteria for selecting the QM-model includes all atoms and groups involved in: (1) bond formation and breaking processes, (2) stabilization of reactant, transition state, and product complexes, and (3) short- and long-range contributions such as electrostatic interactions.<sup>7</sup>

The development of faster and cheaper computers in addition to improving DFT methods has allowed computational chemists to systematically study systems of larger size. Remarkably, it is now possible to study systems containing 250-300 atoms at the QM level compared to less than 50 atoms 20 years ago.<sup>45</sup>

On the other hand, considering only the small reactive region with these methods and ignoring the rest of the system (protein environment in biological systems) has also raised some concerns. First, the environment can impose constraints on the different parts of the system (*e.g.*, steric hindrance and conformational variations).<sup>45-47</sup> Not including it in the calculation might result in large artificial movement of the remaining groups in the chosen-model leading to inaccurate and incorrect descriptions of the (bio)chemical reaction. Second, long-range interactions and electrostatic effects from the environment are omitted in this approach.<sup>45-47</sup> The latter might impact the computed energies. Two simple approximations account for the excluded part of the chemical system: a coordinate-locking scheme and the polarizable continuum model (PCM).<sup>45-47</sup>

To model the steric impact of the environment, certain key atoms, typically those at the truncation's coordinates, are kept locked during the geometry optimizations to prevent unrealistic fluctuations. It is worth pointing out that selecting a model that is too small, leads to a rigid description of the active site and an artificial strain, contributing to the wrong energy surface. On the other hand, a system containing a large number of atoms could face the multiple-minima problem which can be simply solved by applying this

procedure. Hence in mechanistic studies, it is highly recommended to start with a relatively small model and gradually increase the size to reach the optimal-sized QM-cluster model.<sup>45</sup>  
<sup>48</sup> Systematically increasing the size of the QM-cluster model from smaller to a larger one, gives better chemical insight into the group contribution in the studied system.<sup>45, 48-49</sup>

To account for electrostatic and long-range interactions, dielectric cavity techniques are commonly used. Applying this approximation, the environment is treated as a homogenous polarizable medium with different values of dielectric constant ( $\epsilon$ ).<sup>45-47</sup> Despite being arbitrary, a dielectric constant of 4 is found to be an acceptable representation for protein active site.<sup>46, 50</sup> Interestingly, some systematic investigations have shown that solvent influences saturate quickly with the size of the model and the choice of the dielectric constant become less critical as the size of the cluster starts increasing.<sup>46-48, 51</sup> This clearly indicates that the approximation works well when a larger model of the active site is selected; that is when the polarization effects on the reactive center are taken into account explicitly.

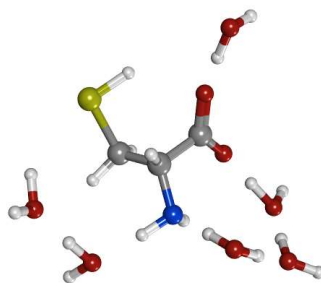
It should be noted that the initial model is typically derived from an X-ray structure or a conformer from molecular dynamics. Either way, the robustness and success of QM models rely upon the quality of the initial structure. The highest resolution crystal structure is usually suggested. As has been computed, at a high resolution of 0.8-1.5 Å the coordinate error is on average < 0.1 Å, while a medium resolution between 1.5 to 2.5 Å increases the coordinate error to 0.3 Å.<sup>52</sup>

In particular, the QM-cluster methodology has been introduced as an extremely valuable protocol to elucidate reaction mechanisms. Indeed, a relatively inexpensive computational cost provides a suitable opportunity to examine and compare various mechanisms with different variables (*i.e.*, solvent, *etc.*) to validate the most feasible one among a dozen of proposed mechanisms.

### 2.7 Solvation Models

As previously mentioned, electrostatics and long-range interactions are consequences of moving from an isolated state (gas phase) to solution. Indeed, physical and chemical properties, reactivities of many (bio)chemical systems, and even reaction results are potentially impacted by the solution and more specifically solvent molecules. Thousands of (bio)chemical reactions, from organic synthesis to enzymatic processes, occur in solution and solvent-free reactions remain a challenge for many chemical experiments. Therefore, accounting for solvent effects in computational studies is crucial to attaining the most reliable results and comprehensive insight into the reaction environment. There are two general approaches to consider the solvent effects in quantum chemistry investigations: the explicit solvation model and the implicit solvation model, each with its own advantages and disadvantages.<sup>3, 53-55</sup>

***Explicit Solvation Model (Supermolecule Model):*** This model consists of an aggregation of a limited number of solvent molecules and solute (see **Figure 2.6**). In other words, solvent molecules are treated explicitly at the desired level of theory in this model. The explicit solvation approach provides detailed information of solvent-solute intermolecular interactions and molecular details of each solvent molecule atomistically through modeling a more realistic physical picture of the system of interest. In contrast, long-range electrostatic interactions cannot be properly modeled by applying this approach. As well, adding more solvent molecules to the system might be problematic due to the multiple-minima problem, additional fluctuation, and increased computational cost.



**Figure 2.6.** Schematic illustration of the explicit solvation model; solute with some solvent molecules included in the calculation.

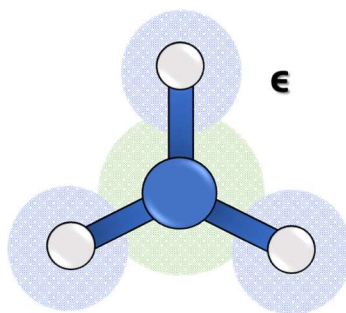
***Implicit Solvation Model (Continuum Model):*** Without doubt, the implicit solvation model is the most common approach in modern computational chemistry. Using this protocol, the polarizable continuum model (PCM), the solvent is treated as a polarizable continuous medium where it can be characterized by its macroscopic properties. In fact, the solute molecule is held in a cavity which is surrounded by the applied continuum. This model is by far a more efficient approach to take into account long-range electrostatic interactions, though not a good model to treat short-range interactions between solute and solvent molecules in an atomistic scale. Computationally, there are a number of implicit solvation models which differ in aspects such as size and shape of the solute cavity, solute charge distribution, and medium description.

Integral equation formalism polarizable continuum model (IEFPCM) is the solvation model employed in this dissertation. Inspired by the pioneering solvation model of Tomasi and coworkers, this model is based on the construction of the solute cavity through a series of the interlocking atomic spheres of their respective van der Waals or ionic radii (see **Figure 2.7**).<sup>56-57</sup>

As mentioned, one of the main differences between the two solvation models is that implicit solvation applies a homogenous medium to treat the solvent implicitly, while the explicit solvation model treats the environment atomistically and gives a more realistic depiction of the entire system. Hybrid models have been developed to combine the



advantages of both models.<sup>3, 53</sup> The simplest hybrid model is a simple combination of the two models; adding explicit solvent and treating the entire system with an implicit model as well. This technique treats the first solvation sphere explicitly to examine intermolecular interactions, while the long-range electrostatic ones are treated *via* a continuum model implicitly.

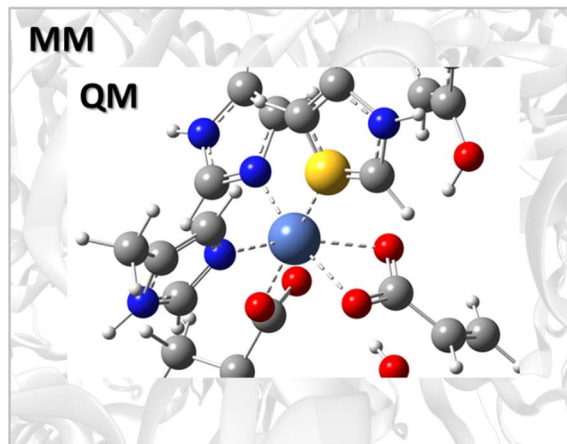


**Figure 2.7.** Schematic illustration of Tomasi's PCM solvation model.

## 2.8 Quantum Mechanics/Molecular Mechanics (QM/MM)

The pioneering work of Warshel and Levitt in 1976 led to the Nobel Prize in chemistry in 2013 for the "development of the combined quantum mechanical/molecular mechanical (QM/MM) technique".<sup>58</sup> Since then, the method has been introduced as one of the most efficient and popular means for studying complex (bio)molecular systems.<sup>6, 59-61</sup> In the QM/MM approach, a region of interest is identified where an electronic structure description is required, such as the region in which chemical processes take place. This layer is treated quantum mechanically while the environment around it is described using a classical molecular mechanics force field (see **Figure 2.8**).

In fact, this methodology ensures that the key chemical processes are modeled with sufficiently accurate and reliable QM methods while it avoids cumbersome computational costs by treating the entire system quantum mechanically.



**Figure 2.8.** Schematic representation of a QM/MM model. The shaded and center regions correspond to the MM and QM layers, respectively.

As mentioned previously, DFT methods are the most preferred level of theory for quantum mechanics calculations, while the AMBER force field is commonly used to treat the MM region for biological contexts

QM/MM methodology differs importantly in terms of some features that should be considered carefully in studying (bio)chemical systems. These aspects include: (1) interaction between the MM and QM regions, (2) computing the total QM/MM potential energy, and (3) crossing the covalent bonds between atoms at the QM/MM boundary.<sup>4, 7, 59, 62</sup>

### 2.8.1 QM and MM Interactions

QM and MM interactions are one of the most critical tasks in the hybrid QM/MM methodology which are addressed carefully in calculations. Generally, interactions between these two subsystems consist of bonded and non-bonded interactions. Bonded interactions include bond stretching, bond bending, and dihedral interactions between two layers while non-bonded interactions involve the electrostatic and van der Waals contributions. Indeed, a very important feature in QM/MM calculations is the way in which electrostatic interactions between QM and MM layers are treated. Considering this type of

interaction, three different protocols can be used in QM/MM computations; (1) mechanical embedding, (2) electrostatic embedding, and (3) polarizable embedding.

Mechanical embedding is the simplest and crudest approach to incorporate the MM layer influence on the QM/MM calculations in the biological systems. In fact, using the mechanical embedding the coupling between two layers is described only at the MM level of theory that could be included in the applied forces field. Electrostatic contributions are neglected and only the van der Waals interaction between the two layers is accounted for. Although this approach is computationally inexpensive, the MM environment cannot induce polarization of the electron density in the QM part, which is the main shortcoming of mechanical treatment.

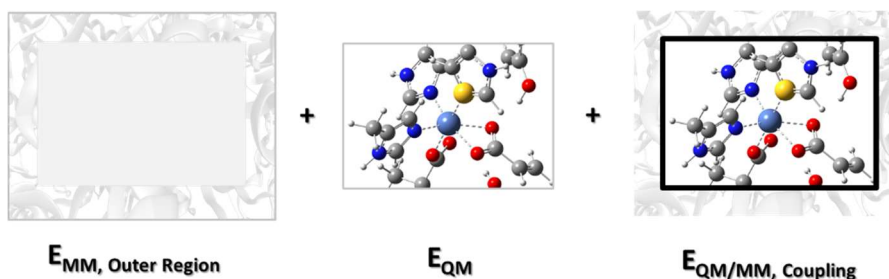
In contrast to mechanical treatment, electrostatic embedding applies a more advanced procedure whereby the QM layer is polarized by the MM charge distribution. In this formalism, the electrostatic interactions between the two layers are taken into account during the electronic wave function computation by the inclusion of a one-electron operator in the QM Hamiltonian. This improvement leads to a more difficult implementation and increased computational cost but more reliable results, particularly in a polar active site. Although not in all cases, the electrostatic formalism might result in an over-polarization of the inner region especially close to the covalent junction between the QM and MM parts while mechanical embedding does not apply any correction scheme and avoids over-polarization of the QM layer by the MM.<sup>63</sup>

In the polarizable embedding formalism, both QM and MM layers mutually polarize each other until self-consistency is obtained in the charge distribution. Thus QM/MM computation can be very demanding and cumbersome for large models. Also using a polarizable scheme needs an accurate polarizable force field which is still one of the holy grails in computational biochemistry.<sup>1</sup>

### 2.8.2 Computing the Total Potential Energy of the QM/MM

The hybrid QM/MM potential energy contains three classes of interactions including the interaction between atoms in the QM layer and the MM layer, separately and interaction of atoms between the two QM and MM layers. The first and second class of interactions is relatively easy to compute, while the interaction between the inner and outer layers is a bit more challenging to describe. To date, several approaches have been proposed to calculate the total potential energy of the QM/MM systems. They can be roughly categorized into two main categories: additive and subtractive schemes.<sup>59, 62</sup>

**Additive QM/MM Scheme:** As shown in **Figure 2.9**, the total QM/MM energy using the additive scheme is the sum of the individual energies of the QM and MM parts, plus a coupling term between QM and MM regions ( $E_{\text{QM/MM, Coupling}}$ ).

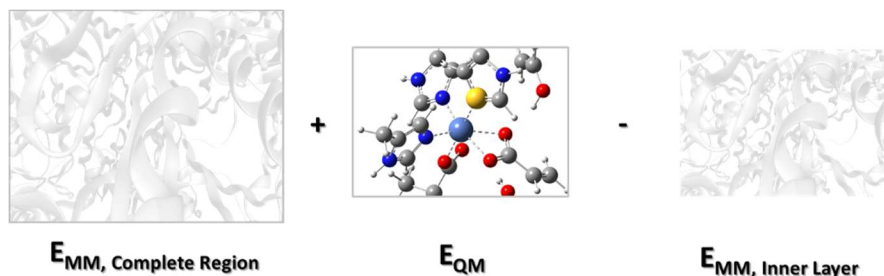


**Figure 2.9.** Schematic representation of an additive coupling within a QM/MM methodology in computing total energy.

Obviously, in the additive scheme, the interaction between two subsystems are treated explicitly. Such interactions include both bonded and non-bonded interactions consisting of van der Waals and electrostatic interactions. Despite being popular in QM/MM implementations, especially in biological systems, the calculation of the QM/MM coupling term might be problematic using the electrostatic embedding and/or when link atoms are present.

**Subtractive QM/MM Scheme:** In the subtractive scheme, the total potential energy of QM/MM can be obtained from calculating three different terms including the energy of the

entire system by MM, the energy of the inner layer both at the QM and MM levels to eliminate the multiple counting (**Figure 2.10**). The main advantage of such a subtractive scheme is that no coupling term between QM and MM regions is required that makes this implementation quite straightforward in the calculations.



**Figure 2.10.** Schematic representation of a subtractive coupling within a QM/MM methodology in computing total energy.

Due to simplicity and implicit cancellation of the artifacts introduced by the link atoms in the subtractive scheme, some methods based on this protocol have been developed by Morokuma and coworkers, for instance, IMOMM and IMOMO.<sup>64-65</sup> One of the most applied subtractive schemes is called “ONIOM”; Own *N*-layer integrated molecular orbital molecular mechanics. This approach can be easily expanded to a combination of multiple layers each treated at a different level of theory such as ONIOM (QM:QM:MM) including two QM and one MM layers. It is worth noting that the second QM layer (the medium region) is treated using a cheaper method such as semiempirical to consider the electronic effects on the first QM part. Nowadays the ONIOM scheme is extensively applied in the investigation of a variety of biomolecular systems.<sup>66-69</sup>

### 2.8.3 QM/MM Boundary Schemes

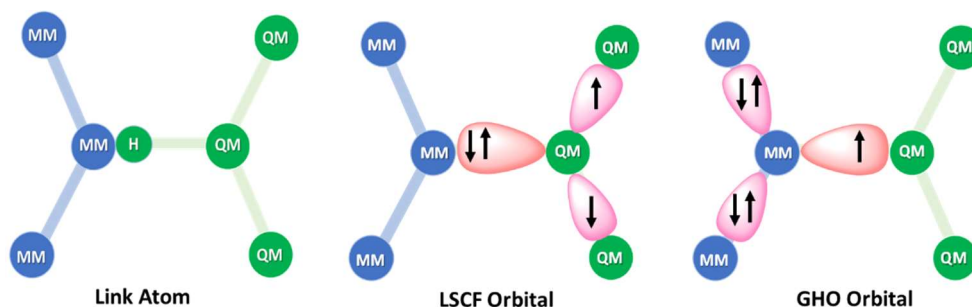
One of the main issues in hybrid QM/MM calculations is related to the QM and MM boundary, where a covalent bond is disrupted and one or more unpaired electrons in the QM calculation results. Two major approaches for dealing with this situation have been

proposed: link atoms and localized orbitals.<sup>62, 70-71</sup> The link atom approach is the most widely used solution in which each free valency is saturated by a so-called link atom. There is no limitation on the type of the link atom, though the hydrogen atom is the most common in computations. Other types of link groups may be used, for instance, methyl groups or pseudohalogens. In principle, each link atom generates three artificial degrees of freedom, however, in practice they are kept fixed along with the partitioned bond between QM and MM atoms in every step of the calculation. Thus the additional degrees of freedom are removed again (**Figure 2.11**).

Using the second solution, a doubly-occupied molecular orbital is replaced by the link atom. Based on this technique, two methods have been developed and used commonly: the localized self-consistent field (LSCF) method and the generalized hybrid orbital (GHO) method (**Figure 2.11**). In the localized self-consistent field method, a set of localized orbitals is placed on the QM atom of the broken bond. The hybrid orbital pointing to the MM atoms is doubly-occupied and during the self-consistent field (SCF) iteration is kept frozen. This is while the other hybrid orbitals are singly-occupied and are optimized along with SCF optimization. In contrast, using the GHO approach, a set of localized orbitals is located on the MM atom of the broken bond. During the SCF, the doubly-occupied orbital pointing to the MM atoms is kept frozen while the singly-occupied orbital pointing to the QM atom is optimized. Additional complications of these methods and requirement for parametrization in the localized orbital procedure has made the link atom method the most accepted protocol to cap the QM atom layer, although no significant changes in results are obtained using the localized orbital compared to the link atom method.<sup>62</sup>

Despite the significant influence of the QM/MM boundary and the inner and outer interactions on the observed results, there is no strict guideline for the region partitioning. There are some general rules to have a well-chosen model such as (1) the boundary should be set between a non-polar bond (*e.g.*, C—C), (2) the boundary should be least 3 bonds away from any bond breaking/forming in the QM layer, (3) cutting conjugated or aromatic

sections should not be done, and (4) the boundary does not allow any charge transfers and hybridization changes between two different layers.



**Figure 2.11.** Illustration of different procedures to cap the QM layer.

## 2.9 Thermodynamic Aspects

**Single Point Energy:** As described earlier, larger basis sets generally lead to more accurate energies despite the high computational costs. A way to calculate such energies and lower the computational cost during the iterative optimization processes is by using larger basis sets to perform single point energy calculation on the optimized geometries obtained at a lower level of theory, according to **Equation 2.4**.

$$\text{Method A/Basis Set A} // \text{Method B/Basis Set B} \quad (\text{Equation 2.4})$$

Where method A and basis set A are the applied level of theory for the single point energy calculation and method B and basis set B is used for geometry optimization and frequency calculation. Also, some energy corrections such as Gibb's free energy and enthalpy energy corrections from frequency calculations can be added to the single point energy to calculate the more accurate and reliable thermodynamic properties.

**Potential Energy Surface:** Mapping the obtained energies of key species in a chemical reaction (*e.g.*, reactant complexes, intermediates, transition structures, and product complexes) generates a potential energy surface (PES). In other words, as a result of the

Born-Oppenheimer approximation, the PES is a function that gives the energy of each molecular structure as a function of nuclear position.<sup>3</sup> Several mechanistic insights such as reaction feasibility, relative energies between molecular structures (*e.g.*, activation energy), structures, properties, and the rate-determining step (RDS) can be accurately elucidated *via* analyzing the PESs.<sup>72-74</sup>

### 2.10 References

1. Houk, K. N.; Liu, F., Holy Grails for Computational Organic Chemistry and Biochemistry. *Acc. Chem. Res.* **2017**, *50*, 539-543.
2. Millikin, R.; Bianco, C. L.; White, C.; Saund, S. S.; Henriquez, S.; Sosa, V.; Akaike, T.; Kumagai, Y.; Soeda, S.; Toscano, J. P.; Lin, J.; Fukuto, J. M., The Chemical Biology of Protein Hydropersulfides: Studies of a Possible Protective Function of Biological Hydropersulfide Generation. *Free Radic. Biol. Med.* **2016**, *97*, 136-147.
3. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*. John Wiley & Sons Ltd.; New York: 2004.
4. Martin-Santamaria, S., *Computational Tools for Chemical Biology*. The Royal Society of Chemistry: London, UK: 2018.
5. Tunon, I.; Moliner, V., *Simulating Enzyme Reactivity*. The Royal Society of Chemistry: UK: 2017.
6. Quesne, M. G.; Borowski, T.; de Visser, S. P., Quantum Mechanics/Molecular Mechanics Modeling of Enzymatic Processes: Caveats and Breakthroughs. *Chem-Eur. J.* **2016**, *22*, 2562-2581.
7. Ahmadi, S.; Barrios Herrera, L.; Chehelamirani, M.; Hostaš, J.; Jalife, S.; Salahub, D. R., Multiscale Modeling of Enzymes: QM-Cluster, QM/MM, and QM/MM/MD: A Tutorial Review. *Int. J. Quantum Chem.* **2018**, *118*, 1-34.



8. Ferreira, P.; Cerqueira, N. M. F. S. A.; Brás, N. F.; Fernandes, P. A.; Ramos, M. J., Parametrization of Molybdenum Cofactors for the AMBER Force Field. *J. Chem. Theory Comput.* **2018**, *14*, 2538-2548.
9. Li, P.; Merz, K. M., Metal Ion Modeling Using Classical Mechanics. *Chem. Rev.* **2017**, *117*, 1564-1686.
10. Neves, R. P. P.; Sousa, S. F.; Fernandes, P. A.; Ramos, M. J., Parameters for Molecular Dynamics Simulations of Manganese-Containing Metalloproteins. *J. Chem. Theory Comput.* **2013**, *9*, 2718-2732.
11. Hu, L.; Ryde, U., Comparison of Methods to Obtain Force-Field Parameters for Metal Sites. *J. Chem. Theory Comput.* **2011**, *7*, 2452-2463.
12. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P., A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999-2012.
13. Foloppe, N.; MacKerell, J., Alexander D., All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *J. Comput. Chem.* **2000**, *21*, 86-104.
14. Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F., A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656-1676.
15. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K., Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781-1802.
16. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26*, 1701-1718.

17. Case, D. A.; Cheatham III, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz Jr., K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668-1688.
18. Christen, M.; Hünenberger, P. H.; Bakowies, D.; Baron, R.; Bürgi, R.; Geerke, D. P.; Heinz, T. N.; Kastenholz, M. A.; Kräutler, V.; Oostenbrink, C.; Peter, C.; Trzesniak, D.; van Gunsteren, W. F., The GROMOS Software for Biomolecular Simulation: GROMOS05. *J. Comput. Chem.* **2005**, *26*, 1719-1751.
19. Brooks, B. R.; Brooks III, C. L.; Mackerell Jr., A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; *et al.* CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545-1614.
20. Koch, W.; Holthausen, M. C., *A Chemist's Guide to Density Functional Theory*. 2nd ed.; Wiley-VCH Verlag GmbH: Weinheim, Germany: 2001.
21. Nikoo, S.; Meister, P. J.; Hayward, J. J.; Gault, J. W., An Assessment of Computational Methods for Calculating Accurate Structures and Energies of Bio-Relevant Polysulfur/Selenium-Containing Compounds. *Molecules* **2018**, *23*, 3323.
22. Baciú, C.; Gault, J. W., An Assessment of Theoretical Methods for the Calculation of Accurate Structures and S–N Bond Dissociation Energies of S-Nitrosothiols (RSNOs). *J. Phys. Chem. A* **2003**, *107*, 9946-9952.
23. Lau, N.; Pluth, M. D., Reactive Sulfur Species (RSS): Persulfides, Polysulfides, Potential, and Problems. *Curr. Opin. Chem. Biol.* **2019**, *49*, 1-8.
24. Sousa, S. F.; Fernandes, P. A.; Ramos, M. J., General Performance of Density Functionals. *J. Phys. Chem. A* **2007**, *111*, 10439-10452.
25. Perdew, J. P.; Ruzsinszky, A.; Tao, J.; Staroverov, V. N.; Scuseria, G. E.; Csonka, G. I., Prescription for the Design and Selection of Density Functional Approximations: More Constraint Satisfaction with Fewer Fits. *J. Chem. Phys.* **2005**, *123*, 1-9.

26. Seifert, G.; Porezag, D.; Frauenheim, T., Calculations of Molecules, Clusters, and Solids with a Simplified LCAO-DFT-LDA Scheme. *Int. J. Quantum Chem.* **1996**, *58*, 185-192.
27. Becke, A. D., Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098-3100.
28. Perdew, J. P.; Wang, Y., Accurate and Simple Analytic Representation of the Electron-Gas Correlation Energy. *Phys. Rev. B* **1992**, *45*, 13244-13249.
29. Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785-789.
30. Becke, A. D., A New Inhomogeneity Parameter in Density-Functional Theory. *J Chem. Phys.* **1998**, *109*, 2092-2098.
31. Becke, A. D., A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372-1377.
32. Mardirossian, N.; Head-Gordon, M., Thirty Years Of Density Functional Theory in Computational Chemistry: an Overview and Extensive Assessment of 200 Density Functionals. *Mol. Phys.* **2017**, *115*, 2315-2372.
33. Cohen, A. J.; Mori-Sánchez, P.; Yang, W., Challenges for Density Functional Theory. *Chem. Rev.* **2012**, *112*, 289-320.
34. Pribram-Jones, A.; Gross, D. A.; Burke, K., DFT: A Theory Full of Holes? *Annu. Rev. Phys. Chem.* **2015**, *66*, 283-304.
35. Zhao, Y.; Truhlar, D. G., Density Functionals with Broad Applicability in Chemistry. *Acc. Chem. Res.* **2008**, *41*, 157-167.
36. Zhao, Y.; Truhlar, D. G., The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215-241.

37. Zhao, Y.; Truhlar, D. G., Exploring the Limit of Accuracy of the Global Hybrid Meta Density Functional for Main-Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2008**, *4*, 1849-1868.
38. Mardirossian, N.; Head-Gordon, M., How Accurate Are the Minnesota Density Functionals for Noncovalent Interactions, Isomerization Energies, Thermochemistry, and Barrier Heights Involving Molecules Composed of Main-Group Elements? *J. Chem. Theory Comput.* **2016**, *12*, 4303-4325.
39. Brea, O.; Daver, H.; Rebek, J.; Himo, F., Mechanism(s) of Thermal Decomposition of N-Nitrosoamides: A Density Functional Theory Study. *Tetrahedron* **2019**, *75*, 929-935.
40. Planas, F.; McLeish, M. J.; Himo, F., Computational Characterization of Enzyme-Bound Thiamin Diphosphate Reveals a Surprisingly Stable Tricyclic State: Implications for Catalysis. *Beilstein J. Org. Chem.* **2019**, *15*, 145-159.
41. Blomberg, M. R. A., How Quantum Chemistry Can Solve Fundamental Problems in Bioenergetics. *Int. J. Quantum Chem.* **2015**, *115*, 1197-1201.
42. Moa, S.; Himo, F., Quantum Chemical Study of Mechanism and Stereoselectivity of Secondary Alcohol Dehydrogenase. *J. Inorg. Biochem.* **2017**, *175*, 259-266.
43. Lind, M. E. S.; Himo, F., Quantum Chemical Modeling of Enantioconvergence in Soluble Epoxide Hydrolase. *ACS Catal.* **2016**, *6*, 8145-8155.
44. Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn, P. E. M., Quantum Chemical Studies of Mechanisms for Metalloenzymes. *Chem. Rev.* **2014**, *114*, 3601-3658.
45. Himo, F., Recent Trends in Quantum Chemical Modeling of Enzymatic Reactions. *J. Am. Chem. Soc.* **2017**, *139*, 6780-6786.
46. Siegbahn, P. E. M.; Himo, F., The quantum Chemical Cluster Approach for Modeling Enzyme Reactions. *WIREs. Comput. Mol. Sci.* **2011**, *1*, 323-336.

47. Chen, S.-L.; Fang, W.-H.; Himo, F., Technical Aspects of Quantum Chemical Modeling of Enzymatic Reactions: the Case of Phosphotriesterase. *Theor. Chem. Acc.* **2008**, *120*, 515-522.
48. Sevastik, R.; Himo, F., Quantum Chemical Modeling of Enzymatic Reactions: The case of 4-oxalocrotonate tautomerase. *Bioorg. Chem.* **2007**, *35*, 444-457.
49. Georgieva, P.; Himo, F., Quantum Chemical Modeling of Enzymatic Reactions: The Case of Histone Lysine Methyltransferase. *J. Comput. Chem.* **2010**, *31*, 1707-1714.
50. Himo, F., Quantum Chemical Modeling of Enzyme Active Sites and Reaction Mechanisms. *Theor. Chem. Acc.* **2006**, *116*, 232-240.
51. Siegbahn, P. E. M.; Himo, F., Recent Developments of the Quantum Chemical Cluster Approach for Modeling Enzyme Reactions. *J. Biol. Inorg. Chem.* **2009**, *14*, 643-651.
52. Liao, R.-Z.; Thiel, W., On the Effect of Varying Constraints in the Quantum Mechanics Only Modeling of Enzymatic Reactions: The Case of Acetylene Hydratase. *J. Phys. Chem. B* **2013**, *117*, 3954-3961.
53. Monard, G.; Rivail, J.-L., Solvent Effects in Quantum Chemistry. In *Handbook of Computational Chemistry*, Leszczynski, J.; Kaczmarek-Kedziera, A.; Puzyn, T.; G. Papadopoulos, M.; Reis, H.; K. Shukla, M., Eds. Springer International Publishing: Cham, 2017; pp 727-739.
54. Cramer, C. J.; Truhlar, D. G., A Universal Approach to Solvation Modeling. *Acc. Chem. Res.* **2008**, *41*, 760-768.
55. Mennucci, B., Polarizable Continuum Model. *WIREs. Comput. Mol. Sci.* **2012**, *2*, 386-404.
56. Cancès, E.; Mennucci, B.; Tomasi, J., A new Integral Equation Formalism for the Polarizable Continuum Model: Theoretical Background and Applications to Isotropic and Anisotropic Dielectrics. *J. Chem. Phys.* **1997**, *107*, 3032-3041.

57. Tomasi, J.; Mennucci, B.; Cammi, R., Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999-3094.
58. Warshel, A.; Levitt, M., Theoretical studies of enzymic reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227-249.
59. Sousa, S. F.; Ribeiro, A. J. M.; Neves, R. P. P.; Brás, N. F.; Cerqueira, N. M. F. S. A.; Fernandes, P. A.; Ramos, M. J., Application of Quantum Mechanics/Molecular Mechanics Methods in the Study of Enzymatic Reaction Mechanisms. *WIREs. Comput. Mol. Sci.* **2017**, *7*, 1-29.
60. Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H.-B.; Ding, L.; Morokuma, K., The ONIOM Method and Its Applications. *Chem. Rev.* **2015**, *115*, 5678-5796.
61. Barbault, F.; Maurel, F., Simulation with Quantum Mechanics/Molecular mechanics for Drug Discovery. *Expert Opin. Drug Dis.* **2015**, *10*, 1047-1057.
62. Groenhof, G., Introduction to QM/MM Simulations. In *Biomolecular Simulations: Methods and Protocols*, Monticelli, L.; Salonen, E., Eds. Humana Press: Totowa, NJ, 2013; pp 43-66.
63. Hu, L.; Söderhjelm, P.; Ryde, U., On the Convergence of QM/MM Energies. *J. Chem. Theory Comput.* **2011**, *7*, 761-777.
64. Maseras, F.; Morokuma, K., IMOMM: A New Integrated Ab Initio + Molecular Mechanics Geometry Optimization Scheme of Equilibrium Structures and Transition States. *J. Comput. Chem.* **1995**, *16*, 1170-1179.
65. Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K., ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels–Alder Reactions and Pt(P(*t*-Bu)<sub>3</sub>)<sub>2</sub> + H<sub>2</sub> Oxidative Addition. *J. Phys. Chem.* **1996**, *100*, 19357-19363.

66. Wilson, K. A.; Fernandes, P. A.; Ramos, M. J.; Wetmore, S. D., Exploring the Identity of the General Base for a DNA Polymerase Catalyzed Reaction Using QM/MM: The Case Study of Human Translesion Synthesis Polymerase  $\eta$ . *ACS Catal.* **2019**, *9*, 2543-2551.
67. Aboelnga, M. M.; Gault, J. W., Roles of the Active Site Zn(II) and Residues in Substrate Discrimination by Threonyl-tRNA Synthetase: An MD and QM/MM Investigation. *J. Phys. Chem. B* **2017**, *121*, 6163-6174.
68. Nogueira, J. J.; Roßbach, S.; Ochsenfeld, C.; González, L., Effect of DNA Environment on Electronically Excited States of Methylene Blue Evaluated by a Three-Layered QM/QM/MM ONIOM Scheme. *J. Chem. Theory Comput.* **2018**, *14*, 4298-4308.
69. Fernandes, H. S.; Ramos, M. J.; Cerqueira, N. M. F. S. A., Catalytic Mechanism of the Serine Hydroxymethyltransferase: A Computational ONIOM QM/MM Study. *ACS Catal.* **2018**, *8*, 10096-10110.
70. Antes, I.; Thiel, W., Adjusted Connection Atoms for Combined Quantum Mechanical and Molecular Mechanical Methods. *J. Phys. Chem. A* **1999**, *103*, 9290-9295.
71. Senn, H. M.; Thiel, W., QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198-1229.
72. Bushnell, E. A. C.; Huang, W.; Gault, J. W., Applications of Potential Energy Surfaces in the Study of Enzymatic Reactions. *Adv. Phys. Chem.* **2012**, *2012*, 15, 1-15.
73. Albaugh, A.; Boateng, H. A.; Bradshaw, R. T.; Demerdash, O. N.; Dziedzic, J.; Mao, Y. Z.; Margul, D. T.; Swails, J.; Zeng, Q.; *et al.* Advanced Potential Energy Surfaces for Molecular Simulation. *J. Phys. Chem. B* **2016**, *120*, 9811-9832.
74. Shchegoleva, L. N.; Beregovaya, I. V., Potential energy Surface as a Key to Understanding the Structure and Properties of Short-Living Radical Ions of Cyclic Organic Molecules. *Int. J. Quantum Chem.* **2016**, *116*, 161-173.

## CHAPTER 3

# **An Assessment of Computational Methods for Calculating Accurate Structures and Energies of Bio-Relevant Polysulfur/Selenium-Containing Compounds**





### 3.1 Introduction

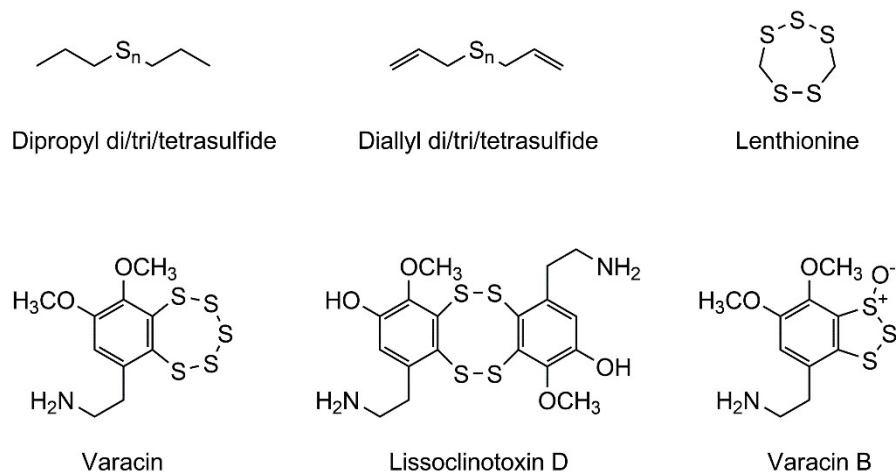
The chalcogens sulfur and selenium have long been known to play key roles in a diverse array of important physiological and biological processes including enzymatic mechanisms,<sup>1-2</sup> signaling,<sup>3-5</sup> and mediation and repair of oxidatively-damaged biomolecules.<sup>6-8</sup> Indeed, in addition to being found within 3 of the 22 proteinogenic amino acids; cysteine (Cys), methionine (Met), and selenocysteine (Sec), they are also found in many essential metabolites (*e.g.*, thiazole). This is due to their ability to possess a broad range of oxidation states as well as bonding environments, and can often undergo reversible redox.<sup>9</sup> For instance, the most abundant antioxidant in animal cells is glutathione (GSH) which can mediate the cellular redox environment through interconversion with its oxidized disulfide form, GSSG.<sup>10</sup>

Recently, it has increasingly been recognized that Reactive Sulfur Species (RSS) comprise a rich and diverse range of physiologically important species (**Scheme 3.1**).<sup>11</sup> Indeed, hydrogen sulfide is now known to be a ubiquitous essential signaling molecule that plays key roles in many physiological and inflammatory processes including blood pressure regulation, cell proliferation and apoptosis, insulin signaling, and neurotransmission.<sup>7, 12-16</sup> However, sulfur can also form strong homonuclear single bonds and as a result can react with other sulfur species (for example, in proteins) to form a variety of polysulfur-containing RSS.<sup>17</sup>

Previously, the presence of per- (HSSH) and polysulfides (*e.g.*,  $\text{H}_2\text{S}_n$   $n=3-7$ ) in biological systems were thought to be experimental artifacts or stores for  $\text{H}_2\text{S}$ .<sup>18</sup> Now, however, they are increasingly proposed or recognized as being biochemically important;<sup>17, 19-22</sup> for instance, some polysulfides have been shown to possess antibiotic or anticancer properties.<sup>23-24</sup> More recently, Cysteinyl-tRNA synthetase (CysRS), an ancient enzyme with a critical role in gene-encoded protein synthesis, has been shown to also catalyze the formation of Cys-derived polysulfides.<sup>21</sup> This further underscores the potential biologically important activity of peptide-hydropersulfides.<sup>25</sup> Meanwhile, Se has been

shown to form an Se-S- intermediate in the selenoproteins thioredoxin reductase and formate dehydrogenase and plays a central role in the enzyme's activation.<sup>26-27</sup> Unfortunately, due to the high reactivity of RSS, specifically per/polysulfides within biological environments, many of their properties and much of their chemistry remains unclear or even unknown.

**Scheme 3.1.** Several examples of naturally occurring polysulfur species.<sup>17, 28</sup>



Computational chemistry has established itself as a key tool for the study of the properties and chemistry of biomolecular systems. Several such studies have been performed on S- and, to a notably lesser extent, Se- or mixed S/Se-containing per- and poly-seleno/sulfides. For example, Brzostowska *et al.* in part used the B3LYP method to examine the intramolecular reactions of the naturally occurring polysulfur-containing pentathiepins (such as varacin, **Scheme 3.1**) that generate S<sub>3</sub> and S<sub>2</sub> transfer units *via* a tetra- or trisulfide anion, respectively.<sup>29</sup> Recently, the high reactivity of several smaller hydropersulfides toward alkyl, alkoxy, peroxy and thiyl radicals was investigated using both experimental and computational (CBS-QB3) methods.<sup>30</sup> It was concluded that such reactions are exothermic by 15-34 kcal mol<sup>-1</sup> due to the low RSS-H bond dissociation enthalpy and high stability of perthiyl radicals.<sup>30</sup> Meanwhile, a computational study has used in part dispersion-corrected B3LYP (B3LYP-D3) to examine the role of the chalcogen

atoms in the mechanism of glutathione peroxidase 4 which involves formation of a -Se-S- species.<sup>31</sup> Bachrach *et al.* used several computational methods including MP2, and B3LYP to examine possible mechanisms for nucleophilic attack at the Se in diselenides and selenosulfides and concluded that attack at Se is kinetically and thermodynamically preferred.<sup>32</sup> Using an ONIOM QM/MM approach wherein the DFT method M06-2X was used for the QM region, Huang *et al.* examined *S*-sulfhydration *via* a persulfide (RSS<sup>-</sup>) intermediate as catalyzed by mercaptopyruvate sulfurtransferase (MST) and obtained reasonable agreement with experiment.<sup>33</sup> In all of these studies, smaller basis sets (*e.g.*, 6-31G(d)) were used to obtain structures upon which they then based their calculations of thermochemical properties. The ability of a computational study to reliably and accurately provide insights into the structures and properties of any biomolecular system usually critically depends on the choice of QM method and basis set. Thus, an essential step towards computationally studying RSS is determining appropriate methodologies.

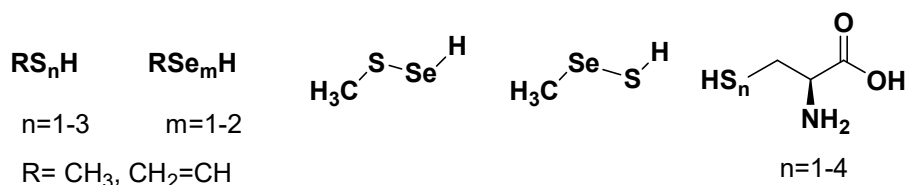
In this present study, the ability of a range of computational, in particular density functional theory (DFT) methods<sup>34</sup>, to provide reliable and accurate structures and thermochemical properties of biologically relevant poly-sulfur/selenium-containing compounds has been assessed. More specifically, the DFT methods B3LYP, B3PW91,  $\omega$ B97XD, M06-2X, and M08-HX were applied to a systematic series of biologically relevant RX<sub>*n*</sub>(H) (X=S/Se, *n*=1-3, R=CH<sub>3</sub>, CH<sub>2</sub>=CH, and cysteine) species. As well as their structures, a variety of key thermochemical properties including their X-X (chalcogen-chalcogen) bond dissociation enthalpies, hydrogen affinities, and gas phase basicities were examined and benchmarked.

### 3.2 Computational Methods

All calculations were performed using the Gaussian 09<sup>35</sup> and Gaussian 16<sup>36</sup> suites of programs. Optimized geometries for a systematic series of biologically relevant chemical models, shown in **Scheme 3.2**, were obtained using a variety of conventional wavefunction

and density functional theory methods applied in conjunction with a range of Pople basis sets from 6-31G(d) to 6-311++G(3df,3pd).

**Scheme 3.2.** Schematic illustration of the species considered in this study.



Specifically, the hybrid DFT methods B3LYP and B3PW91, comprised of Becke's three parameter exchange functional<sup>37</sup> (B3) in combination with the Lee, Yang, and Parr correlation functional<sup>38</sup> (LYP) or the Perdew and Wang functional<sup>39</sup> (PW91), were assessed. B3LYP is commonly applied in the treatment of biomolecular systems while B3PW91 has been used previously for systems containing sulfur and selenium.<sup>40-41</sup> In addition, two *meta*-GGA functionals were assessed; M06-2X<sup>42</sup>, a commonly employed functional in the study of enzymatic catalysis, and the more recently developed related functional M08-HX.<sup>43</sup> For this functional, optimized structures were obtained using Gaussian 16 whereas all other structures were obtained with Gaussian 09. Furthermore, the double-hybrid range-corrected functional  $\omega$ B97XD<sup>44</sup> was also evaluated to further examine any effects of dispersion correction on geometric or energetic parameters. Geometries optimized at these levels of theory were compared to those obtained using the *ab initio* MP2 (see **Appendix A3**) and QCISD methods. As DFT methods are the preferred choice for biochemical systems due to their computational cost and reliability, the MP2 results are only given in **Appendix A3.1** for our baseline geometry assessment of  $CH_3SSH^-$  and  $CH_3SeSeH^-$ . The Minnesota family of functionals (M06-2X and M08HX) are more empirical and explicitly contain dispersion correction while B3LYP only contains 3 empirical parameters and has no explicit correction for dispersion. Likewise,  $\omega$ B97XD contains dispersion corrections.

All optimized structures were confirmed to be minima by harmonic vibrational frequency calculations performed at the same level of theory. These were also used to determine zero-point vibrational energy (ZPVE) and enthalpy corrections for subsequent calculations of a range of their thermochemical properties including proton affinities ( $^{298.15\text{K}}\text{PA}_{\text{A}^-} = -\Delta\text{H} = -(\text{H}_{\text{AH}} - \text{H}_{\text{A}^-} - \text{H}_{\text{H}^+})$ ), gas-phase basicities ( $^{298.15\text{K}}\text{GPB}_{\text{A}^-} = -\Delta\text{G} = -(-\text{PA} - \text{T}(\text{S}_{\text{AH}} - \text{S}_{\text{A}^-} - \text{S}_{\text{H}^+}))$ ), and hydrogen affinities ( $^{298.15\text{K}}\text{HA}_{\text{A}} = -\Delta\text{H} = -(\text{H}_{\text{AH}} - \text{H}_{\text{A}^\bullet} - \text{H}_{\text{H}^\bullet})$ ).<sup>45-46</sup> In addition, homolytic bond dissociation enthalpies were calculated for production of the  $\cdot\text{SH}$  or  $\cdot\text{SeH}$  radicals ( $^{298.15\text{K}}\text{BDE}_{\text{AX-XH}} = \Delta\text{H} = \text{H}_{\text{AX}^\bullet} + \text{H}_{\text{XH}^\bullet} - \text{H}_{\text{AXXH}}$ ).

### 3.3 Results and Discussion

#### 3.3.1 Structural Assessment of $\text{CH}_3\text{XXH}$ and $\text{CH}_3\text{XX}^-$ (X=S, Se)

We began by using the broadest variety of DFT methods (B3LYP, B3PW91,  $\omega\text{B97XD}$ , M06-2X, M08-HX) and range of basis sets (6-31G(d) to 6-311++G(3df,3pd)) used in this study to obtain optimized structures for  $\text{CH}_3\text{XXH}$  and  $\text{CH}_3\text{XX}^-$  (X=S, Se). These are the smallest homoatomic persulf/selenides considered in this present study. Due to a paucity of corresponding experimental data, benchmark optimized structures were obtained at the QCISD/6-311+G(2df,p) level of theory. For simplicity, only the key C–X, X–X, and X–H (X=S, Se) distances are discussed herein and are summarized in **Table 3.1** and MP2 data is shown in **Appendix A3.1**.

##### 3.3.2.1 Method Sensitivity to Basis Set Changes

All five DFT methods showed similar overall sensitivities to changes in the basis set from 6-31G(d) to 6-311++G(3df,3pd). For example, the smallest variations were observed for the C–X and X–H (X=S, Se) bond distances; for any given method and chemical system (be it neutral or anionic) they varied by  $\leq 0.020$  Å. The only exceptions occurred for the C–Se bond in  $\text{CH}_3\text{SeSe}^-$  when using M06-2X method which varied overall by 0.027 Å, and the Se–H bond using MP2 which varied by 0.025 Å (**Appendix A3.1**). Notably, for

any DFT method and chemical system the C—S bond distances showed greater variation than that of C—Se bonds. The only exceptions to this trend occurred for the C—X bonds in the  $\text{CH}_3\text{XX}^-$  anions when using the B3LYP or M06-2X methods. However, the converse trend was observed for the X—H bonds; optimized Se—H bond lengths are more sensitive to basis set choice than S—H bonds (**Tables 3.1** and **Appendix A3.1**).

For all chemical systems (both neutral and anionic), regardless of the choice of DFT method, the largest variations upon changing basis set were observed in their X—X bonds. For  $\text{CH}_3\text{SSH}$  and  $\text{CH}_3\text{SS}^-$  it was found that the S—S bond distances varied by 0.032 – 0.043 and 0.044 – 0.051 Å, respectively. Meanwhile, in the corresponding  $\text{CH}_3\text{SeSeH}$  and  $\text{CH}_3\text{SeSe}^-$  species the Se—Se bond distances varied by 0.024 – 0.029 and 0.023 – 0.029 Å, respectively, except for using MP2 where the variation exceeded the trend for DFT data (**Appendix A3.1**). That is, except for Se—H bonds, those bonds involving sulfur (*i.e.*, C—X and X—X) in  $\text{CH}_3\text{XXH}$  and  $\text{CH}_3\text{XX}^-$  (X=S, Se) are most sensitive to the choice of basis set.

**Table 3.1.** Selected optimized bond lengths in ångström (Å) for CH<sub>3</sub>XXH and CH<sub>3</sub>XX<sup>-</sup> (X=S, Se).

Method	Basis set	CH <sub>3</sub> SSH			CH <sub>3</sub> SS <sup>-</sup>		CH <sub>3</sub> SeSeH			CH <sub>3</sub> SeSe <sup>-</sup>	
		C-S	S-S	S-H	C-S	S-S	C-Se	Se-Se	Se-H	C-Se	Se-Se
<b>B3LYP</b>	6-31G(d)	1.834	2.091	1.357	1.837	2.117	1.977	2.337	1.495	1.985	2.366
	6-311G(d)	1.830	2.103	1.357	1.831	2.134	1.979	2.362	1.488	1.987	2.391
	6-311G(d,p)	1.831	2.105	1.354	1.833	2.134	1.978	2.364	1.478	1.987	2.391
	6-311+G(d,p)	1.830	2.103	1.354	1.832	2.128	1.978	2.364	1.478	1.987	2.388
	6-311G(2d,p)	1.830	2.088	1.349	1.831	2.115	1.975	2.366	1.476	1.984	2.395
	6-311G(df,p)	1.829	2.093	1.354	1.831	2.120	1.973	2.351	1.479	1.981	2.376
	6-311+G(2df,p)	1.825	2.072	1.351	1.827	2.092	1.971	2.350	1.477	1.979	2.373
	6-311++G(3df,3pd)	1.820	2.064	1.349	1.822	2.083	1.971	2.351	1.476	1.978	2.373
<b>B3PW91</b>	6-31G(d)	1.821	2.072	1.354	1.824	2.096	1.961	2.315	1.491	1.960	2.336
	6-311G(d)	1.817	2.081	1.356	1.819	2.110	1.963	2.338	1.486	1.963	2.361
	6-311G(d,p)	1.818	2.083	1.353	1.820	2.110	1.962	2.340	1.476	1.962	2.360
	6-311+G(d,p)	1.818	2.083	1.353	1.821	2.105	1.962	2.339	1.476	1.962	2.358
	6-311G(2d,p)	1.817	2.067	1.349	1.819	2.091	1.959	2.342	1.474	1.959	2.363
	6-311G(df,p)	1.816	2.072	1.353	1.818	2.097	1.957	2.327	1.477	1.958	2.345
	6-311+G(2df,p)	1.812	2.053	1.351	1.814	2.071	1.955	2.327	1.475	1.955	2.342
	6-311++G(3df,3pd)	1.808	2.046	1.349	1.810	2.063	1.955	2.327	1.475	1.955	2.342
<b>ωB97XD</b>	6-31G(d)	1.819	2.070	1.349	1.821	2.095	1.954	2.309	1.486	1.968	2.342
	6-311G(d)	1.816	2.079	1.350	1.816	2.110	1.957	2.331	1.480	1.971	2.365
	6-311G(d,p)	1.816	2.081	1.348	1.817	2.110	1.956	2.333	1.470	1.970	2.365
	6-311+G(d,p)	1.816	2.080	1.348	1.817	2.105	1.956	2.333	1.470	1.970	2.362
	6-311G(2d,p)	1.816	2.067	1.344	1.816	2.092	1.953	2.334	1.469	1.967	2.369
	6-311G(df,p)	1.815	2.070	1.348	1.815	2.099	1.951	2.320	1.472	1.965	2.351
	6-311+G(2df,p)	1.810	2.053	1.346	1.811	2.074	1.950	2.319	1.470	1.963	2.348
	6-311++G(3df,3pd)	1.806	2.046	1.344	1.807	2.066	1.949	2.320	1.469	1.962	2.347
<b>M06-2X</b>	6-31G(d)	1.819	2.069	1.348	1.819	2.096	1.954	2.306	1.486	1.985	2.366
	6-311G(d)	1.816	2.078	1.348	1.816	2.110	1.960	2.330	1.480	1.964	2.358
	6-311G(d,p)	1.817	2.079	1.346	1.817	2.110	1.959	2.331	1.472	1.964	2.358
	6-311+G(d,p)	1.816	2.089	1.346	1.816	2.105	1.959	2.331	1.472	1.964	2.355
	6-311G(2d,p)	1.817	2.066	1.342	1.816	2.093	1.956	2.334	1.471	1.961	2.362
	6-311G(df,p)	1.814	2.068	1.347	1.816	2.097	1.954	2.320	1.473	1.960	2.345
	6-311+G(2df,p)	1.811	2.052	1.345	1.812	2.073	1.953	2.319	1.472	1.958	2.340
	6-311++G(3df,3pd)	1.808	2.046	1.342	1.808	2.065	1.953	2.319	1.472	1.958	2.341
<b>M08-HX</b>	6-31G(d)	1.817	2.068	1.350	1.817	2.094	1.955	2.305	1.488	1.958	2.330
	6-311G(d)	1.816	2.078	1.345	1.815	2.108	1.961	2.327	1.482	1.964	2.352
	6-311G(d,p)	1.816	2.079	1.348	1.815	2.108	1.960	2.328	1.474	1.963	2.352
	6-311+G(d,p)	1.816	2.079	1.349	1.815	2.103	1.960	2.328	1.474	1.964	2.348
	6-311G(2d,p)	1.814	2.067	1.344	1.814	2.093	1.957	2.331	1.473	1.960	2.357
	6-311G(df,p)	1.814	2.069	1.349	1.814	2.095	1.954	2.318	1.475	1.959	2.341
	6-311+G(2df,p)	1.809	2.053	1.347	1.810	2.072	1.953	2.317	1.474	1.957	2.335
	6-311++G(3df,3pd)	1.807	2.047	1.344	1.808	2.063	1.953	2.318	1.472	1.957	2.336
<b>QCISD</b>	<b>6-311+G(2df,p)</b>	<b>1.815</b>	<b>2.065</b>	<b>1.347</b>	<b>1.816</b>	<b>2.088</b>	<b>1.965</b>	<b>2.333</b>	<b>1.474</b>	<b>1.961</b>	<b>2.359</b>

## 3.3.2.2 Effects of Increasing Basis Set Size

As can be seen in **Table 3.1**, for all DFT methods considered and for both neutral  $\text{CH}_3\text{XXH}$  and anionic  $\text{CH}_3\text{XX}^-$  ( $\text{X}=\text{S}, \text{Se}$ ), similar trends were generally observed upon increasing the basis set from double- to triple-zeta, and then subsequently by inclusion of diffuse and polarization functions.

For instance, increasing the basis set from 6-31G(d) to 6-311++G(3df,3pd) generally caused a systematic shortening in the  $\text{X}-\text{H}$  and  $\text{C}-\text{X}$  bonds. Notably, one does not need to increase the basis set significantly in order to get reasonable agreement with the corresponding values obtained at the QCISD/6-311+G(2df,p) level of theory. In fact, upon changing the basis set from 6-31G(d) to 6-311G(d) (*i.e.*, double- to triple-zeta) resulted in their lengths differing from the benchmark values by  $\leq 0.015 \text{ \AA}$ . Further increases in basis set size by inclusion of diffuse (*i.e.*, 6-311G(d,p) to 6-311+G(d,p)) or *f*- and/or *d*-polarization functions on heavy atoms or hydrogen resulted in only minor individual decreases.

In contrast, modifying the basis set showed quite different trends for the  $\text{X}-\text{X}$  ( $\text{X}=\text{S}, \text{Se}$ ) bonds. For instance, for all methods assessed, improving the basis set from 6-31G(d) to 6-311G(d) resulted in an increase in their optimized length for all chemical systems of up to  $0.025 \text{ \AA}$ . The only exception occurred for the M06-2X method applied to  $\text{CH}_3\text{SeSe}^-$  for which the  $\text{Se}-\text{Se}$  bond shortened slightly by  $0.008 \text{ \AA}$ . Notably, in  $\text{CH}_3\text{SSH}$  and  $\text{CH}_3\text{SS}^-$  the observed ranges of bond lengthening were  $0.009 - 0.015$  and  $0.014 - 0.027 \text{ \AA}$  respectively, with the largest increases observed when using the B3LYP method. That is to say, the anionic persulfide is more sensitive to basis set changes than the neutral hydropersulfide. In contrast, for  $\text{CH}_3\text{SeSeH}$  and  $\text{CH}_3\text{SeSe}^-$  the observed increases were quite similar lying in the range of  $0.017 - 0.025 \text{ \AA}$  (except where noted above). The further inclusion of *p*-functions on hydrogen (*i.e.*, 6-311G(d) to 6-311G(d,p)) or diffuse functions on heavy atoms (*i.e.*, 6-311G(d,p) to 6-311+G(d,p)) had negligible effect on the  $\text{X}-\text{X}$  bond lengths in both the neutral  $\text{CH}_3\text{XXH}$  and anionic  $\text{CH}_3\text{XX}^-$  ( $\text{X}=\text{S}, \text{Se}$ ) systems.



In general, more significant changes in the X—X lengths were observed upon inclusion of either a second set of *d*- or a set of *f*-functions on heavy atoms (*i.e.*, 6-311G(d,p) to 6-311G(2d,p) or 6-311G(df,p)). Specifically, for both these basis set changes the S—S bonds in CH<sub>3</sub>SSH and CH<sub>3</sub>SS<sup>−</sup> shortened by 0.010 to 0.017 Å, except for MP2 where the distance increases for the 6-311G(2d,p) basis set. In contrast, for the corresponding selenium containing systems, the inclusion of a second set of *d*-functions on heavy atoms marginally lengthened Se—Se bonds by  $\leq 0.004$  Å, while the inclusion of a set of *f*-functions on heavy atoms shortened the Se—Se bonds by 0.010 – 0.017 Å. Combining or adding further diffuse and polarization functions by use of the 6-311+G(2df,p) or 6-311++G(3df,3pd) basis sets respectively, resulted in all DFT methods except B3LYP giving X—X (X=S, Se) bond lengths that were markedly shorter than their corresponding benchmark values. In the case of B3LYP (which overestimates the length of these bonds) increasing the basis set to 6-311+G(2df,p) or 6-311++G(3df,3pd) was in fact required in order to get good agreement with the benchmark values.

Importantly, for all DFT methods that were considered (with the exception of B3LYP), the 6-311G(2d,p) and 6-311G(df,p) basis sets gave optimized C—X, X—X, and X—H distances for CH<sub>3</sub>XXH and CH<sub>3</sub>XX<sup>−</sup> (X=S) that were in closest general agreement with their corresponding benchmark values. Meanwhile, for the corresponding Se analogues, the best performing basis sets were generally 6-311G(d), 6-311G(d,p), 6-311+G(d,p), and 6-311G(2d,p). It is also noted that the M08-HX method does not offer much if any improvement over M06-2X. In fact, when Se is in the system, it slightly underestimates the bond lengths. As such, subsequent tables showing optimized parameters will only include results obtained using the B3PW91, ωB97XD and M06-2X methods in combination with the identified preferred basis sets. It should be noted that, for completeness, the corresponding values for the other methods are included in the **Appendix A3**. Since MP2 trends were mostly similar to DFT, we do not include results obtained with this method.

### 3.3.2 Effect of Conjugation: $\text{CH}_2=\text{CHXXH}$ and $\text{CH}_2=\text{CHXX}^-$ (X=S, Se)

These model systems were examined to gain insights into the influence of conjugation on the neutral and anionic persulfide and perselenide groups. Based on the results obtained for the  $\text{CH}_3\text{XX}^-/\text{H}$  systems, optimized structures were obtained using only the B3PW91,  $\omega\text{B97XD}$ , and M06-2X methods in combination with the 6-311+G(d,p), 6-311G(2d,p), and 6-311G(df,p) basis sets. The most significant changes were observed in their C—X and X—X bonds, hence only these optimized values are shown in **Table 3.2**. **Appendices A3.2** and **A3.3** include the full set of optimized parameters with all DFT functionals and the X—H bonds.

**Table 3.2.** Selected optimized bond lengths (ångström, Å) for  $\text{CH}_2=\text{CHXXH}$  and  $\text{CH}_2=\text{CHXX}^-$  (X=S, Se).

Method	Basis set	$\text{CH}_2=\text{CHSSH}$		$\text{CH}_2=\text{CHSS}^-$		$\text{CH}_2=\text{CHSeSeH}$		$\text{CH}_2=\text{CHSeSe}^-$	
		C—S	S—S	C—S	S—S	C—Se	Se—Se	C—Se	Se—Se
<b>B3PW91</b>	6-311+G(d,p)	1.768	2.088	1.745	2.098	1.913	2.347	1.899	2.359
	6-311G(2d,p)	1.767	2.074	1.747	2.084	1.910	2.350	1.898	2.364
	6-311G(df,p)	1.768	2.077	1.748	2.090	1.908	2.334	1.894	2.346
<b><math>\omega\text{B97XD}</math></b>	6-311+G(d,p)	1.771	2.084	1.748	2.101	1.913	2.338	1.898	2.357
	6-311G(2d,p)	1.770	2.072	1.750 <sup>a</sup>	2.088 <sup>a</sup>	1.910	2.339	1.897	2.361
	6-311G(df,p)	1.770	2.074	1.751	2.096	1.908	2.325	1.895	2.343
<b>M06-2X</b>	6-311+G(d,p)	1.770	2.084	1.747	2.102	1.914	2.337	1.901	2.354
	6-311G(2d,p)	1.770	2.071	1.751	2.090	1.911	2.340	1.900	2.359
	6-311G(df,p)	1.770	2.072	1.750	2.095	1.909	2.326	1.897	2.343
<b>QCISD</b>	6-311+G(2df,p)	<b>1.771</b>	<b>2.069</b>	<b>1.750</b>	<b>2.086</b>	<b>1.914</b>	<b>2.329</b>	<b>1.902</b>	<b>2.359</b>

<sup>a</sup> Had one negligible imaginary frequency.

In general, upon changing basis sets and methods similar trends were observed as for the  $\text{CH}_3\text{XX}^-/\text{H}$  systems, though with some key differences. For instance, for X=S, the closest agreement of the optimized bond lengths with the benchmark values was obtained for the 6-311G(df,p) and 6-311G(2d,p) basis sets. In contrast, for X=Se the X—X bond is more sensitive to the choice of basis set. In particular, for  $\text{CH}_2\text{CHSeSe}^-$  the best agreement with the benchmark values are obtained using the 6-311+G(d,p) or 6-311G(2d,p) basis sets; inclusion of *f*-functions results in too short an Se—Se bond. Meanwhile, for neutral

CH<sub>2</sub>CHSeSeH it is essential to include *f*-functions in order to obtain good agreement with the corresponding benchmark values.

Comparison of the optimized bond lengths in CH<sub>3</sub>XX<sup>-</sup>/H and CH<sub>2</sub>CHXX<sup>-</sup>/H (X=S, Se) shows that for all species, at all levels of theory, the optimized C—X distance in CH<sub>2</sub>CHXX<sup>-</sup>/H is 0.04 – 0.05 Å shorter than in the corresponding CH<sub>3</sub>XX<sup>-</sup>/H analogue (*cf.* **Table 3.1**). This is also observed when structures are optimized at the M06-2X/aug-cc-pVTZ level of theory (data not shown). In contrast, the X—X bond lengths in CH<sub>2</sub>CHXX<sup>-</sup>/H (X=S, Se) are all within ±0.01 Å of their optimized values in the corresponding CH<sub>3</sub>XX<sup>-</sup>/H analogue. When X=S the X—X bond in the anion is 0.017 Å longer in the conjugated species compared to 0.023 Å in the alkyl terminated species. It is noted that this again indicates that deprotonation of RXXH (X=S, Se) causes the X—X bond to lengthen, although conjugation lessens the bond lengthening upon going to the anion.

### 3.3.3 Effects of Mixed Sulfur/Selenium: CH<sub>3</sub>XYH (X=S, Se; Y=Se, S)

We then considered mixed sulfur/selenide species for which selected optimized parameters are given in **Table 3.3**. Full optimized parameters for all DFT methods and basis sets are shown in **Appendix A3.4**.

For all four mixed species the optimized lengths of the C—X bond (X=S, Se) are all within 0.01 Å of those obtained for their analogous persulfide or perselenide (*i.e.*, CH<sub>3</sub>XXH (X=S, Se)) at the same level of theory (*cf.* **Table 3.1**). Thus, similar method and basis set trends were also observed and do not require further detailed discussion. However, it should be noted that while the X—Y bond lengths in the neutral species CH<sub>3</sub>SSeH and CH<sub>3</sub>SeSH are close to each other, the bond is consistently predicted to be marginally shorter (< 0.01 Å) in CH<sub>3</sub>SSeH. The changes observed in the CH<sub>3</sub>X—Y bond length upon deprotonation (*i.e.*, CH<sub>3</sub>XYH to CH<sub>3</sub>XY<sup>-</sup>), depends on whether S or Se is the terminal atom. When S is the terminal atom, upon deprotonation of the thiol group the Se—S bond lengthens marginally by < 0.01 Å. In contrast, when Se is the terminal atom deprotonation

causes the S—Se bond to lengthen by  $\geq 0.04$  Å. The larger impact of deprotonating a selenol *versus* thiol group is also seen in the species shown in **Tables 3.1** and **3.2**, though to a smaller extent. We found that when the sulfur atom is in the middle of the chain, it has a larger Mulliken charge compared to selenium ( $-0.23$  vs.  $-0.13$ ) in the anionic species. By comparison, the charge of the terminal chalcogen is the same in both species. Thus, repulsive interactions cause the bond length to increase when sulfur is in the center of the chain.

**Table 3.3.** Selected optimized bond lengths (ångström, Å) for  $\text{CH}_3\text{XYH}$  and  $\text{CH}_3\text{XY}^-$  ( $\text{X}=\text{S}, \text{Se}$ ;  $\text{Y}=\text{Se}, \text{S}$ ).

Method	Basis set	$\text{CH}_3\text{SSeH}$		$\text{CH}_3\text{SSe}^-$		$\text{CH}_3\text{SeSH}$		$\text{CH}_3\text{SeS}^-$	
		C—S	S—Se	C—S	S—Se	C—Se	Se—S	C—Se	Se—S
<b>B3PW91</b>	6-311+G(d,p)	1.821	2.214	1.820	2.256	1.959	2.218	1.971	2.221
	6-311G(2d,p)	1.819	2.204	1.819	2.244	1.957	2.208	1.968	2.220
	6-311G(df,p)	1.819	2.199	1.819	2.241	1.953	2.204	1.964	2.212
<b><math>\omega</math>B97XD</b>	6-311+G(d,p)	1.820	2.208	1.817	2.252	1.953	2.213	1.962	2.218
	6-311G(2d,p)	1.818	2.199	1.817	2.242	1.951	2.204	1.959	2.216
	6-311G(df,p)	1.818	2.195	1.817	2.239	1.948	2.200	1.957	2.209
<b>M06-2X</b>	6-311+G(d,p)	1.819	2.208	1.817	2.250	1.956	2.212	1.964	2.220
	6-311G(2d,p)	1.817	2.200	1.816	2.239	1.954	2.203	1.960	2.221
	6-311G(df,p)	1.817	2.194	1.816	2.237	1.951	2.198	1.959	2.212
<b>QCISD</b>	6-311+G(2df,p)	<b>1.819</b>	<b>2.198</b>	<b>1.816</b>	<b>2.239</b>	<b>1.952</b>	<b>2.202</b>	<b>1.961</b>	<b>2.207</b>

### 3.3.4 Extending a Sulfide: $\text{CH}_3\text{SSSH}$ and $\text{CH}_2\text{CHSSSH}$

Polysulfides, but to-date not polyselenides, have been suggested to be potentially biochemically important. Thus, for completeness we considered the effect of extending the persulfide chain. Specifically, the simplest alkyl and conjugated-containing trisulfides  $\text{CH}_3\text{SSSH}$  and  $\text{CH}_2\text{CHSSSH}$  were examined. The C—S<sub>1</sub> bond in all species had very similar optimized lengths, for all methods and basis sets used, to those obtained for the analogous alkyl persulfides (*cf.* **Table 3.1**); the observed shortened C—X bond in the conjugated neutral and anionic persulfides (**Table 3.2**) was not observed in  $\text{CH}_2\text{CHSSSH}$ , further evidence that it is due to delocalization across the persulf/selenide and conjugated R-group. Thus, in **Table 3.4** only the optimized lengths of the S—S bonds are given.

As can be seen in **Table 3.4**, the B3PW91 method gives the worst agreement with the benchmark values, especially for the deprotonated (anionic) species with errors of up to 0.07 Å. Furthermore, for any DFT method, the largest errors in the optimized  $S_i-S_j$  bond lengths are observed upon use of the 6-311+G(d,p) basis set. The best agreement is instead obtained using the M06-2X and  $\omega$ B97XD methods, the former performing slightly better in conjunction with the 6-311G(2d,p) basis set. Indeed, their errors lie in the ranges of 0.000 – 0.011 Å and – 0.003 – 0.009 Å, respectively.

**Table 3.4.** Selected optimized bond lengths (Å) for RSSSH (R=CH<sub>3</sub>, CH<sub>2</sub>=CH).

Method	Basis set	CH <sub>3</sub> SSSH		CH <sub>3</sub> SSS <sup>−</sup>		CH <sub>2</sub> CHSSSH		CH <sub>2</sub> CHSSS <sup>−</sup>	
		S <sub>1</sub> –S <sub>2</sub>	S <sub>2</sub> –S <sub>3</sub>	S <sub>1</sub> –S <sub>2</sub>	S <sub>2</sub> –S <sub>3</sub>	S <sub>1</sub> –S <sub>2</sub>	S <sub>2</sub> –S <sub>3</sub>	S <sub>1</sub> –S <sub>2</sub>	S <sub>2</sub> –S <sub>3</sub>
<b>B3PW91</b>	6-311+G(d,p)	2.069	2.102	2.147	2.067	2.082	2.094	2.170	2.052
	6-311G(2d,p)	2.057	2.086	2.120	2.058	2.071	2.079	2.136	2.047
	6-311G(df,p)	2.060	2.091	2.130	2.063	2.073	2.084	2.149	2.050
<b><math>\omega</math>B97XD</b>	6-311+G(d,p)	2.068	2.093	2.124	2.073	2.078	2.086	2.134	2.062
	6-311G(2d,p)	2.057	2.080	2.105	2.063	2.070	2.074	2.115	2.054
	6-311G(df,p)	2.059	2.084	2.111	2.070	2.071	2.077	2.123	2.060
<b>M06-2X</b>	6-311+G(d,p)	2.066	2.090	2.121	2.074	2.079	2.083	2.143	2.063
	6-311G(2d,p)	2.056	2.078	2.103	2.065	2.070	2.071	2.118	2.055
	6-311G(df,p)	2.057	2.080	2.107	2.069	2.070	2.073	2.123	2.060
<b>QCISD</b>	6-311+G(2df,p)	<b>2.056</b>	<b>2.078</b>	<b>2.096</b>	<b>2.065</b>	<b>2.066</b>	<b>2.073</b>	<b>2.107</b>	<b>2.057</b>

Notably, in CH<sub>3</sub>SSSH and CH<sub>2</sub>CHSSSH, the CS<sub>1</sub>—S<sub>2</sub> bond is predicted to be shorter than the S<sub>2</sub>—S<sub>3</sub>H bond by > 0.02 and < 0.01 Å respectively. However, upon deprotonation of the terminal thiol group in each, *i.e.*, formation of CH<sub>3</sub>SSS<sup>−</sup> and CH<sub>2</sub>CHSSS<sup>−</sup>, the CS<sub>1</sub>—S<sub>2</sub> bond lengthens significantly by ~0.05 Å, from ~2.06 and 2.07 Å in CH<sub>3</sub>SSSH and CH<sub>2</sub>CHSSSH respectively, to approximately 2.1 and 2.12 Å. In contrast, the S<sub>2</sub>—S<sub>3</sub> bond in CH<sub>3</sub>SSS<sup>−</sup> and CH<sub>2</sub>CHSSS<sup>−</sup> has shortened by 0.01–0.02 Å. Mulliken charges on S<sub>1</sub> were found to decrease more than they do on S<sub>2</sub> upon deprotonation in the benchmark calculation, for both alkyl and conjugated polysulfur species. However, the difference between the two was small compared to the large charge on S<sub>3</sub>; it becomes much more negative upon deprotonation. This indicates there is some degree of charge delocalization along the sulfur chain (**Appendix A3.6**).

### 3.3.5 Obtaining Reliable and Accurate Thermochemistry for CH<sub>3</sub>SSH and CH<sub>2</sub>CHSSH

Two of the most common goals when applying computational methods to the study of chemical problems are obtaining reliable and accurate optimized structures and thermochemical data. For biochemical or related problems, common reactions often require knowledge of proton affinities (PAs), gas phase basicities (GPBs), hydrogen affinities (HAs) and bond dissociation enthalpies (BDEs). Furthermore, given the size of the systems often encountered there is simultaneously considerable interest in identifying a DFT-based approach for calculating such properties. Thus, having assessed the use of DFT methods for the accurate optimization of structures, we also assessed the ability of the DFT methods B3LYP, B3PW91,  $\omega$ B97XD, M06-2X, M08-HX, in combination with a range of basis sets to provide reliable and accurate biochemically-relevant thermochemical data.

Given the poor performance of B3LYP and minimal improvement of M08-HX over M06-2X in obtaining reliable structures, *vide supra*, here we only report the performance of B3PW91,  $\omega$ B97XD, and M06-2X, unless otherwise noted. Furthermore, we have focussed on reporting basis sets that for such systems have been previously used (*e.g.*, 6-31G(d)), shown herein to be most consistently reliable (*i.e.*, 6-311G(2d,p)), or often used for calculating accurate thermochemistry (*i.e.*, 6-311+G(2df,p) and 6-311++G(3df,3pd)). The results obtained are summarized in **Table 3.5**, although the data for all functionals and basis sets that were studied are shown in **Appendices A3.7 - A3.10**.

The benchmark values were again obtained using the QCISD/6-311+G(2df,p) level of theory. The thermochemical values obtained for the conjugated persulfides compared to the alkyl persulfides reflects the trends observed structurally. For instance, the BDE(S–S) for CH<sub>2</sub>CHSSH is slightly higher by 5.7 kJ mol<sup>–1</sup> than that of CH<sub>3</sub>SSH, while the PA and GPB of CH<sub>2</sub>CHSS<sup>–</sup> are 23.6 and 24.4 kJ mol<sup>–1</sup> lower, respectively, than those of CH<sub>3</sub>SS<sup>–</sup>. This is due to delocalization across the persulfide and CH<sub>2</sub>CH group in CH<sub>2</sub>CHSSH/<sup>–</sup>.

Meanwhile, the hydrogen affinity of  $\text{CH}_2\text{CHSS}^\bullet$  is predicted to be only marginally higher than that of  $\text{CH}_3\text{SS}^\bullet$  by  $0.8 \text{ kJ mol}^{-1}$ . Notably, we did see spin contamination in the QCISD calculations with a spin of 0.88. This may be the reason for lower energies calculated by QCISD compared to DFT methods where spin contamination was much lower (0.77).

A DFT-based model was determined to be accurate if it gave values within  $10 \text{ kJ mol}^{-1}$ , generally held to be experimental accuracy, of the benchmark values. From **Table 3.5** it can be seen that not all methods or basis set choices were reliable, nor did all thermochemical properties exhibit the same method/basis set requirement. For instance, for accurate determination of the BDE(S—S) of both  $\text{CH}_3\text{SSH}$  and  $\text{CH}_2\text{CHSSH}$ , the best performance was observed for B3PW91 and  $\omega\text{B97XD}$  in combination with the 6-31G(d) basis set. The former method slightly underestimating compared to the benchmark value while the latter overestimated slightly. The M06-2X method only gave an accurate BDE(S—S) for  $\text{CH}_2\text{CHSSH}$  and again when using the 6-31G(d) basis set. In general, the use of basis sets larger than 6-31G(d) gave BDE(S—S) values that are markedly higher ( $14 - 36 \text{ kJ mol}^{-1}$ ) than those obtained at the QCISD/6-311+G(2df,p) level of theory.

For the PA and GPB of the  $\text{CH}_3\text{SS}^-$  and  $\text{CH}_2\text{CHSS}^-$  anions the M06-2X method again only gives accurate values when used in combination with the 6-31G(d) basis set. In contrast, the B3PW91 and  $\omega\text{B97XD}$  give accurate values for all the basis sets considered. However, the triple-zeta basis sets (6-311G(2d,p), 6-311+G(2df,p), and 6-311++G(3df,3pd)) gave best agreement with calculated values within  $5 \text{ kJ mol}^{-1}$  of their corresponding benchmark value (see **Table 3.5**).

**Table 3.5.** Homolytic S–S bond dissociation enthalpy (BDE) of RSSH, proton affinity (PA) and gas-phase basicity (GPB) of  $\text{RSS}^-$ , and hydrogen affinity (HA) of  $\text{RSS}^*$  ( $\text{R}=\text{CH}_3$ ,  $\text{CH}_2\text{CH}$ ). All energies calculated at 298.15K and in  $\text{kJ mol}^{-1}$ .

Method	Basis Set	$\text{CH}_3\text{SSH}$	$\text{CH}_3\text{SS}^-$		$\text{CH}_3\text{SS}^*$
		BDE(S–S)	PA	GPB	HA
<b>B3PW91</b>	6-31G(d)	235.4	1455.2	1424.8	274.6
	6-311G(2d,p)	250.3	1447.7	1417.3	283.8
	6-311+G(2df,p)	254.7	1445.8	1415.4	279.5
	6-311++G(3df,3pd)	255.8	1448.7	1418.4	281.6
<b><math>\omega</math>B97XD</b>	6-31G(d)	244.1	1455.0	1425.1	284.0
	6-311G(2d,p)	258.5	1447.3	1417.5	292.0
	6-311+G(2df,p)	264.6	1445.5	1415.8	288.0
	6-311++G(3df,3pd)	265.6	1449.2	1419.6	289.8
<b>M06-2X</b>	6-31G(d)	252.0	1440.0	1411.4	285.4
	6-311G(2d,p)	264.9	1430.0	1401.6	294.4
	6-311+G(2df,p)	270.3	1428.5	1400.1	291.0
	6-311++G(3df,3pd)	272.3	1432.3	1404.2	293.9
<b>QCISD</b>	6-311+G(2df,p)	<b>236.3</b>	<b>1445.4</b>	<b>1415.1</b>	<b>282.7</b>
		$\text{CH}_2\text{CHSSH}$	$\text{CH}_2\text{CHSS}^-$		$\text{CH}_2\text{CHSS}^*$
		BDE(S–S)	PA	GPB	HA
<b>B3PW91</b>	6-31G(d)	238.0	1427.6	1396.5	272.8
	6-311G(2d,p)	253.0	1423.1	1391.5	282.1
	6-311+G(2df,p)	255.7	1417.7	1386.8	278.4
	6-311++G(3df,3pd)	256.8	1420.5	1389.7	280.1
<b><math>\omega</math>B97XD</b>	6-31G(d)	244.1	1429.8	1397.3	283.7
	6-311G(2d,p)	258.5	1422.4 <sup>a</sup>	1395.9 <sup>a</sup>	291.5
	6-311+G(2df,p)	262.7	1420.2	1382.0	288.3
	6-311++G(3df,3pd)	263.5	1423.6	1388.7	289.7
<b>M06-2X</b>	6-31G(d)	249.8	1414.5	1381.1	283.7
	6-311G(2d,p)	263.0	1408.2	1374.4	292.8
	6-311+G(2df,p)	267.4	1403.4	1366.4	290.2
	6-311++G(3df,3pd)	269.1	1406.7	1368.1	292.6
<b>QCISD</b>	6-311+G(2df,p)	<b>242.0</b>	<b>1421.8</b>	<b>1390.7</b>	<b>283.5</b>

<sup>a</sup> One of the required species for determining this value had one negligible imaginary frequency.

The calculated values of the hydrogen affinity of  $\text{CH}_3\text{SS}^*$  and  $\text{CH}_2\text{CHSS}^*$  follow almost the same method and basis set trends and accuracy as that observed for the PA and GPB of  $\text{CH}_3\text{SS}^-$  and  $\text{CH}_2\text{CHSS}^-$ . Namely, the B3PW91 and  $\omega$ B97XD methods in conjunction with any of the basis sets considered give calculated values within  $\pm 10 \text{ kJ mol}^{-1}$ . The only exception occurs when at the B3PW91/6-31G(d) level of theory for  $\text{CH}_2\text{CHSS}^*$ , which



gives a HA value  $10.7 \text{ kJ mol}^{-1}$  lower than the corresponding QCISD/6-311+G(2df,p) benchmark value (see **Table 3.5**). Meanwhile, the M06-2X method is inconsistent; for  $\text{CH}_3\text{SS}^*$  only the 6-31G(d) and 6-311+G(2df,p) basis sets give values within  $10 \text{ kJ mol}^{-1}$  of the benchmark values while for  $\text{CH}_2\text{CHSS}^*$  it gives good agreement for all basis sets considered herein.

Several overall trends are suggested in this examination of the performance of the DFT methods B3PW91,  $\omega$ B97XD, and M06-2X, in conjunction with a range of basis sets, for the noted important thermochemical properties. In particular, the M06-2X functional is the least consistent and usually gives values that differ from the benchmark values by more than  $10 \text{ kJ mol}^{-1}$ . Furthermore, for all DFT functionals the values obtained using the 6-311+G(2df,p) basis set are within  $4 \text{ kJ mol}^{-1}$  of the corresponding values obtained using the much larger and more expensive 6-311++G(3df,3pd) basis set. Hence, for the remainder of this report, for simplicity, only thermochemical values obtained using the B3PW91 and  $\omega$ B97XD functionals in combination with basis sets no larger than 6-311+G(2df,p) are discussed, unless otherwise noted (see **Appendices A3.7 - A3.10** for the complete data sets).

### 3.3.6 Thermochemistry of Selenium-Containing Species

As noted above in the structural assessment, for selenium-containing species considered herein; that is  $\text{CH}_3\text{SeSe}^*/\text{H}$ ,  $\text{CH}_2\text{CHSeSe}^*/\text{H}$ ,  $\text{CH}_3\text{SSe}^*/\text{H}$ , and  $\text{CH}_3\text{SeS}^*/\text{H}$ , the smallest consistently reliable basis set was 6-311+G(d,p), though with exceptions as noted. Hence, for these species we have limited our discussion herein to results obtained using B3PW91 and  $\omega$ B97XD in conjunction with the 6-311+G(d,p) and 6-311+G(2df,p) basis sets. The results are shown in **Table 3.6**.

As for the analogous purely sulfur-containing species (*cf.* **Table 3.5**), the calculated PA's and GPB's of all species obtained using B3PW91 or  $\omega$ B97XD with either basis set choice give values within  $10 \text{ kJ mol}^{-1}$  of the corresponding benchmark values. A similar

consistency is observed for the calculated HA's, though with some exceptions. In particular, at the  $\omega$ B97XD/6-311+G(d,p) level of theory the calculated HA's of  $\text{CH}_2\text{CHSeSe}^\bullet$  and  $\text{CH}_3\text{SSe}^\bullet$  are 11.2 and 12.4  $\text{kJ mol}^{-1}$ , respectively, higher than their corresponding QCISD/6-311+G(2df,p) values. Again, as seen in **Table 3.5**, the calculated RX—YH BDE's of all species are generally markedly overestimated by 12.7 – 26.3  $\text{kJ mol}^{-1}$  using either DFT method and basis set. Only three values fall within the desired 10  $\text{kJ mol}^{-1}$  error margin and all occur for the mixed chalcogen species; all using the 6-311+G(d,p) basis set.

The thermochemical values provided in **Tables 3.5** and **3.6** show several key differences between persulfides and perselenides and the mixed chalcogens. These are most clearly and simply illustrated by examination of the calculated benchmark values. In particular, increasing the number of Se atoms in an RX—YH bond reduces its BDE as shown by comparing those of  $\text{CH}_3\text{SSH}$  (236.3  $\text{kJ mol}^{-1}$ ),  $\text{CH}_3\text{SSeH}$  (218.7  $\text{kJ mol}^{-1}$ ),  $\text{CH}_3\text{SeSH}$  (223.2  $\text{kJ mol}^{-1}$ ), and  $\text{CH}_3\text{SeSeH}$  (202.7  $\text{kJ mol}^{-1}$ ). In addition, a conjugated group adjacent to the RX—XH group increases its BDE slightly by 3-6  $\text{kJ mol}^{-1}$ . Meanwhile, the HA's are reasonably consistent and depend primarily on whether the formal radical terminal is a sulfur or selenium. For the former, all values lie within the range 282.7  $\text{kJ mol}^{-1}$  ( $\text{CH}_3\text{SS}^\bullet$ ) to 286.1  $\text{kJ mol}^{-1}$  ( $\text{CH}_3\text{SeS}^\bullet$ ), while the latter are notably lower, between 265.9  $\text{kJ mol}^{-1}$  ( $\text{CH}_2\text{CHSeSe}^\bullet$ ) and 267.3  $\text{kJ mol}^{-1}$  ( $\text{CH}_3\text{SeSe}^\bullet$  and  $\text{CH}_3\text{SSe}^\bullet$ ).

Comparison of the PA's and GPB's of these perselenides and mixed per-sulf/selenides with their analogous persulfides (*cf.* **Table 3.5**), shows several interesting trends. It is noted that for simplicity, as the observed trends were the same for all the DFT methods, the values discussed here refer to those obtained at the benchmark level of theory. Firstly, systematically increasing the occurrence of Se in a per-chalcogenide group decreases their PA's and GPB values. For instance, the PA's of  $\text{CH}_3\text{SS}^-$ ,  $\text{CH}_3\text{SeS}^-$ ,  $\text{CH}_3\text{SSe}^-$ , and  $\text{CH}_3\text{SeSe}^-$  are 1445.4  $\text{kJ mol}^{-1}$ , 1438.9  $\text{kJ mol}^{-1}$ , 1411.3  $\text{kJ mol}^{-1}$ , and 1404.3  $\text{kJ mol}^{-1}$ . Simultaneously, their GPB values decrease from 1415.1  $\text{kJ mol}^{-1}$ , to 1408.6  $\text{kJ mol}^{-1}$  and

1381.1 kJ mol<sup>-1</sup>, to 1374.0 kJ mol<sup>-1</sup>. In addition, replacing CH<sub>3</sub>- with CH<sub>2</sub>CH- decreases the PA and GPB values of the pure persulfides and perselenides by 24-25 and 17-18 kJ mol<sup>-1</sup>, respectively. This perhaps reflects a larger delocalization when a conjugated group is adjacent and that the effect is less for the selenides.

**Table 3.6.** Calculated homolytic Se–Se bond dissociation enthalpies (BDE) of RSeSeH, proton affinity (PA) and gas-phase basicities (GPB) of RSeSe<sup>-</sup>, and Se–H homolytic bond dissociation enthalpies (HA) of RSeSe<sup>•</sup> (R=CH<sub>3</sub>, CH<sub>2</sub>CH). All energies in kJ mol<sup>-1</sup>.

Method	Basis Set	CH <sub>3</sub> SeSeH	CH <sub>3</sub> SeSe <sup>-</sup>		CH <sub>3</sub> SeSe <sup>•</sup>
		BDE(Se–Se)	PA	GPB	HA
<b>B3PW91</b>	6-311+G(d,p)	219.2	1409.0	1378.7	270.5
	6-311+G(2df,p)	222.9	1411.4	1381.1	266.0
<b>ωB97XD</b>	6-311+G(d,p)	218.0	1410.8	1380.7	276.5
	6-311+G(2df,p)	222.6	1413.3	1383.4	270.7
<b>QCISD</b>	6-311+G(2df,p)	<b>204.1</b>	<b>1404.3</b>	<b>1374.0</b>	<b>267.3</b>
		CH <sub>2</sub> CHSeSeH	CH <sub>2</sub> CHSeSe <sup>-</sup>		CH <sub>2</sub> CHSeSe <sup>•</sup>
		BDE(Se–Se)	PA	GPB	HA
<b>B3PW91</b>	6-311+G(d,p)	222.1	1388.6	1357.4	270.2
	6-311+G(2df,p)	225.3	1390.7	1359.7	265.4
<b>ωB97XD</b>	6-311+G(d,p)	221.3	1391.4	1359.6	277.1
	6-311+G(2df,p)	225.3	1393.8	1361.1	271.0
<b>QCISD</b>	6-311+G(2df,p)	<b>208.3</b>	<b>1386.7</b>	<b>1355.7</b>	<b>265.9</b>
		CH <sub>3</sub> SSeH	CH <sub>3</sub> SSe <sup>-</sup>		CH <sub>3</sub> SSe <sup>•</sup>
		BDE(S–Se)	PA	GPB	HA
<b>B3PW91</b>	6-311+G(d,p)	221.7	1413.2	1382.8	273.8
	6-311+G(2df,p)	233.3	1415.7	1385.5	265.3
<b>ωB97XD</b>	6-311+G(d,p)	219.6	1414.0	1384.2	279.9
	6-311+G(2df,p)	232.5	1417.2	1387.5	270.9
<b>QCISD</b>	6-311+G(2df,p)	<b>218.7</b>	<b>1411.3</b>	<b>1381.1</b>	<b>267.5</b>
		CH <sub>3</sub> SeSH	CH <sub>3</sub> SeS <sup>-</sup>		CH <sub>3</sub> SeS <sup>•</sup>
		BDE(Se–S)	PA	GPB	HA
<b>B3PW91</b>	6-311+G(d,p)	229.4	1435.0	1404.8	287.1
	6-311+G(2df,p)	241.6	1441.7	1411.4	283.6
<b>ωB97XD</b>	6-311+G(d,p)	231.4	1435.6	1405.8	295.6
	6-311+G(2df,p)	245.0	1442.2	1412.4	290.8
<b>QCISD</b>	6-311+G(2df,p)	<b>224.5</b>	<b>1438.9</b>	<b>1408.6</b>	<b>286.1</b>

### 3.3.7 Extending the Persulfides to Trisulfides

We also examined the effect of extending the persulfide group by an additional sulfur to a trisulfide; specifically, we considered key bio-relevant thermochemistry of  $\text{CH}_3\text{SSS}^\bullet/\text{H}$  and  $\text{CH}_2\text{CHSSS}^\bullet/\text{H}$ . Based on the previous trends observed, however, thermochemical values were only obtained using the B3PW91 and  $\omega\text{B97XD}$  DFT methods in combination with the 6-311+G(2df,p) basis set, and again at the QCISD/6-311+G(2df,p) benchmark level of theory. As can be seen in **Table 3.7**, both DFT methods give thermochemical results in generally good agreement with the benchmark values with B3PW91 slightly preferred, although the differences in average errors are relatively small.

Comparison of the calculated benchmark values of the trisulfides with those of the corresponding persulfides (*cf.* **Table 3.6**) shows that the BDE of  $\text{RS}_1\text{S}_2\text{—S}_3\text{H}$  (*i.e.*, the BDE of the terminal  $\text{S}_2\text{—S}_3$  bond), where  $\text{R}=\text{CH}_3\text{—}$  and  $\text{CH}_2\text{CH—}$ , decreases significantly by 53.3 and 56.3  $\text{kJ mol}^{-1}$ , respectively. Similarly, their calculated PAs and GPBs decrease markedly by 31-33 and 18-19  $\text{kJ mol}^{-1}$  for  $\text{R}=\text{CH}_3\text{—}$  and  $\text{CH}_2\text{CH—}$ , respectively. As a result of these changes the  $\text{RSS—SH}$  BDE, and PAs and GPBs of  $\text{RSSS}^\bullet$  all lie within a narrower range ( $\leq 13 \text{ kJ mol}^{-1}$ ) of each other. This perhaps reflects in part a decrease in the influence of the R group on the increasingly removed  $\text{S—SH}$  bond. In contrast, the hydrogen affinities of  $\text{CH}_3\text{SSS}^\bullet$  and  $\text{CH}_2\text{CHSSS}^\bullet$  are 6.7 and 11.5  $\text{kJ mol}^{-1}$  higher than that of their corresponding persulfide analogues with values of 289.4 and 295.0  $\text{kJ mol}^{-1}$ , respectively.

**Table 3.7.** Homolytic RS–SH bond dissociation enthalpy (BDE) of RSSSH, proton affinity (PA) and gas-phase basicity (GPB) of RSSS<sup>−</sup>, and hydrogen affinity (HA) of RSSS<sup>•</sup> (R=CH<sub>3</sub>, CH<sub>2</sub>CH). All energies calculated at 298.15K and in kJ mol<sup>−1</sup>.

Method	Basis Set	CH <sub>3</sub> SSSH	CH <sub>3</sub> SSS <sup>−</sup>		CH <sub>3</sub> SSS <sup>•</sup>
		BDE(RSS—SH)	PA	GPB	HA
<b>B3PW91</b>	6-311+G(2df,p)	193.8	1413.9	1382.4	287.9
<b>ωB97XD</b>	6-311+G(2df,p)	200.4	1415.0	1385.1	297.7
<b>QCISD</b>	6-311+G(2df,p)	<b>183.0</b>	<b>1413.9</b>	<b>1384.1</b>	<b>289.4</b>
		CH <sub>2</sub> CHSSSH	CH <sub>2</sub> CHSSS <sup>−</sup>		CH <sub>2</sub> CHSSS <sup>•</sup>
		BDE(RSS—SH)	PA	GPB	HA
<b>B3PW91</b>	6-311+G(2df,p)	192.8	1399.5	1367.0	290.5
<b>ωB97XD</b>	6-311+G(2df,p)	201.0	1403.0	1370.9	302.0
<b>QCISD</b>	6-311+G(2df,p)	<b>185.7</b>	<b>1403.7</b>	<b>1371.2</b>	<b>295.0</b>

### 3.3.8 Cysteine-Derived Polysulfides: CysSS<sub>n</sub>H (*n*=1-3)

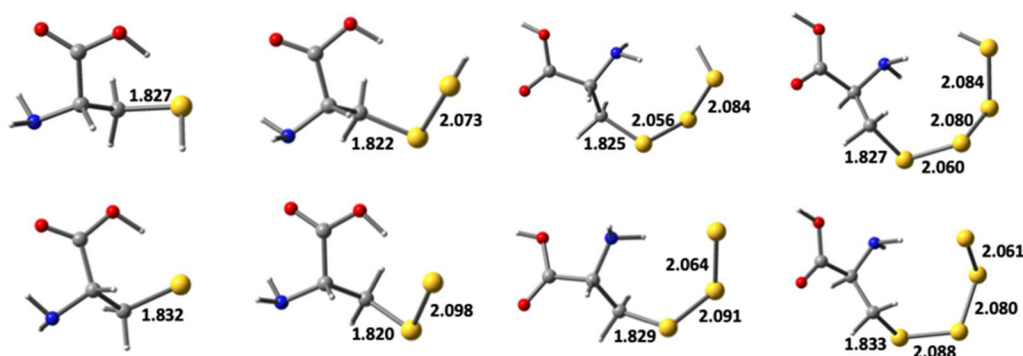
Within biological systems, as noted above, per- and polysulfide derivatives of cysteine play important roles. Hence, we also considered the structures and thermochemistry of such species using the ωB97XD/6-311G(2d,p) level of theory. This level was selected as it emerged in our earlier study as the most able to provide reliable structures and thermochemistry for related model systems (see above). Selected parameters of optimized structures obtained for CysSS<sub>n</sub>H (*n*=0-3) and CysSS<sub>n</sub><sup>−</sup> (*n*=0-3) are provided in **Table 3.8** and representative optimized structures are shown in **Figure 3.1**, at this level of theory.

As can be seen, for both the neutral and deprotonated derivatives, the *r*(C–S) bond length decreases slightly upon forming a perthiol/sulfide group, but then gradually lengthens, essentially returning to its length in cysteine, as one goes to the corresponding trithiol/sulfide and tetrathiol/sulfide derivatives. It is also noted that for both CysSSH and CysSS<sup>−</sup>, the C—S and S—S bond lengths are in close agreement with the corresponding values obtained at the same level of theory for CH<sub>3</sub>SSH (1.816 and 2.067 Å) and CH<sub>3</sub>SS<sup>−</sup> (1.816 and 2.092 Å), respectively (*cf.* **Table 3.1**). In addition, for the neutral polysulfides

the S—S bond in the chain that is farthest removed from the alkyl group has the longest length, with the S—S bonds being sequentially shorter the closer they are to the alkyl group. In contrast, for the corresponding anionic deprotonated series ( $\text{CysSS}_n^-$ ,  $n=0-3$ ), the opposite trend is observed; the S—S bond in the chain farthest from the alkyl group is shortest, and they get longer the closer they are to the alkyl group. In the latter species this trend may reflect a diminishing effect of the negative charge on the terminal sulfur atom the further removed the bond.

**Table 3.8.** Selected optimized bond lengths (ångström, Å) obtained at the  $\omega\text{B97XD/6-311G(2d,p)}$  level of theory for  $\text{CysSS}_n\text{H}$  ( $n=0-3$ ) and  $\text{CysSS}_n^-$  ( $n=0-3$ ).

Molecule	C—S <sub>1</sub>	S <sub>1</sub> —S <sub>2</sub>	S <sub>2</sub> —S <sub>3</sub>	S <sub>3</sub> —S <sub>4</sub>
$\text{CysSH}$	1.827			
$\text{CysS-SH}$	1.822	2.073		
$\text{CysS-S-SH}$	1.825	2.056	2.084	
$\text{CysS-S-S-SH}$	1.827	2.060	2.080	2.084
$\text{CysS}^-$	1.832			
$\text{CysS-S}^-$	1.820	2.098		
$\text{CysS-S-S}^-$	1.829	2.091	2.064	
$\text{CysS-S-S-S}^-$	1.833	2.088	2.080	2.061



**Figure 3.1.** Optimized structures (with C—S and S—S bond length given in ångström (Å)) obtained at the  $\omega\text{B97XD/6-311G(2d,p)}$  level of theory for neutral cysteine and its polysulfides ( $\text{CysSS}_n\text{H}$ ,  $n=1-3$ ) derivatives (top row), and their corresponding anionic deprotonated derivatives ( $\text{CysSS}_n^-$ ,  $n=0-3$ ; bottom row).

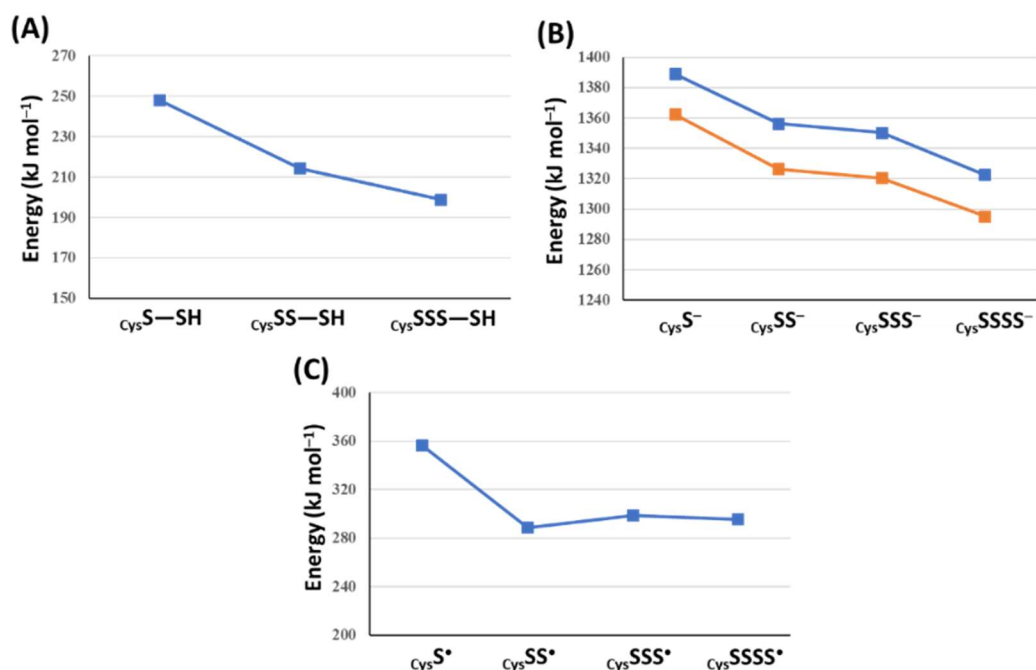
## 3.3.8.1 Thermochemistry of Cysteine-Derived Polysulfides

As for the other species considered above, we calculated the  $RS_n$ —SH homolytic BDE's for  $_{Cys}SS_nH$  ( $n=1-3$ ), the PA's and GPB's of the  $_{Cys}SS_n^-$  ( $n=0-3$ ) series of derivatives, and the hydrogen affinities (HA's) of the  $_{Cys}SS_n^\bullet$  ( $n=0-3$ ) series of species. However, all values were again only obtained using the chosen  $\omega B97XD/6-311G(2d,p)$  level of theory. The results obtained are shown in **Figure 3.2A-C**.

The calculated  $RS_n$ —SH ( $n=1-3$ ) homolytic bond dissociation enthalpies, that is, the BDE of the terminal S—SH bond in the perthiol chain decreases by 33.7 kJ mol<sup>-1</sup> from 248.1 to 214.4, upon going from the perthiol ( $n=1$ ) to trithiol ( $n=2$ ). It then decreases a further 15.4 kJ mol<sup>-1</sup> to 199.0 kJ mol<sup>-1</sup> upon increasing to the chain further to the tetrathiol ( $n=3$ ); **Figure 3.2A**. This trend suggests that while the strength of the terminal S—SH bond does weaken as the chain is lengthened, it approaches a limiting value below 199.0 kJ mol<sup>-1</sup>, although possibly not too much lower than that. In contrast, as seen in **Figure 3.2C**, the hydrogen affinities of the  $_{Cys}SS_n^\bullet$  ( $n=0-3$ ) species decrease significantly by 67.8 kJ mol<sup>-1</sup> upon going from a cysteinyl thiyl radical ( $n=0$ ; 356.5 kJ mol<sup>-1</sup>) to the perthiyl radical ( $n=1$ ; 288.7 kJ mol<sup>-1</sup>). However, extending the chain further to  $n=2$  or  $n=3$  results in a slight increase in the HA's to 298.7 kJ mol<sup>-1</sup> and 295.6 kJ mol<sup>-1</sup>, respectively. This suggests that at least for HA's of the polysulfur radicals, they are reasonably constant for the perthiyl and beyond.

**Figure 3.2B** shows that as the polysulfide chain is extended from the cysteine thiolate to  $_{Cys}SSSS^-$  both the calculated PA's and GPB's decrease. However, while there are significant decreases of 33.0 (PA) and 35.9 (GPB) kJ mol<sup>-1</sup> going from the thiolate ( $_{Cys}S^-$ ) to perthiolate ( $_{Cys}SS^-$ ), this does not appear to continue upon extending the chain further. Indeed, extending the chain to  $_{Cys}SSS^-$  results in only comparatively small further decreases of 5.8 and 5.9 kJ mol<sup>-1</sup> to 1350.5 and 1320.6 kJ mol<sup>-1</sup>. But, upon extending the chain to  $_{Cys}SSSS^-$ , larger decreases in both the PA and GPB are again observed; they decrease by 27.9 and 25.4 kJ mol<sup>-1</sup> to 1322.6 and 1295.2 kJ mol<sup>-1</sup>. This may in part reflect that in the

cysteinyl polysulfide derivatives, weak intramolecular hydrogen bonding was observed in some optimized structures between the terminal  $S^-$  center and the cysteine's amino group. This interaction would also help decrease PA and GPB values. It should also be noted that comparison of the calculated PA's and GPB's of the cysteine-derived sulfides  $CysSS^-$  and  $CysSSS^-$  with that of their smaller corresponding analogs  $CH_3SS^-$  (*cf.* **Table 3.5**) and  $CH_3SSS^-$  (*cf.* **Table 3.7**) shows that those of the former two species are markedly lower by 60-90  $\text{kJ mol}^{-1}$ . Thus, overall, these results suggest that in biological systems, extending the polysulfide chain will increase its likelihood of being deprotonated and that hydrogen bonding may help to stabilize such anions.



**Figure 3.2.** Calculated (at the  $\omega$ B97XD/6-311G(2d,p) level of theory) values of (A)  $RS_n-SH$  BDE's for  $CysSS_nH$  ( $n=1-3$ ), (B) PA's (blue) and GPB's (orange) of  $CysSS_n^-$  ( $n=0-3$ ), and (C) HA's of  $CysSS_n^*$  ( $n=0-3$ ). All energies are in  $\text{kJ mol}^{-1}$ .

### 3.4 Conclusions

The reliability and accuracy of several commonly used DFT functionals (*e.g.*, B3LYP, B3PW91,  $\omega$ B97XD, M06-2X, M08-HX) as well as MP2 was assessed for a systematic



series of bio-relevant polysulfur/selenium-containing systems. In particular, optimized structures and thermodynamic properties of a range of  $RX_n(H)$  ( $X=S, Se, R=CH_3, CH_2=CH$ , and cysteine,  $n=1-4$ ) were examined with a variety of Pople basis sets of increasing size. We offer the following conclusions from this detailed study:

1. Evaluation of the bond lengths in  $CH_3XXH^-$  ( $X=S, Se$ , cf. **Table 3.1**) showed that the S—S bond is the most sensitive to changes in basis set. The smallest basis set used (6-31G(d)) frequently resulted in optimized geometries that were very similar to the benchmark given by QCISD/6-311+G(2df,p), although the preferred basis sets were 6-311G(2d,p) and 6-311G(df,p) due to the higher sensitivity of sulfur to basis set choice.
2. We also saw that M08-HX did not offer much improvement over M06-2X and in some cases was detrimental. The best functionals for geometry optimization were found to be B3PW91,  $\omega$ B97XD and M06-2X.
3. In the conjugated system  $CH_2CHXXH$ , the C—X bond length decreases due to more delocalized electron density.
4. In mixed chalcogen species, the location of the chalcogen atom has a significant effect on the X—Y bond length. The  $RSe-S^-$  bond is slightly longer than the  $RS-Se^-$  bond since sulfur takes on more negative charge. Deprotonation of polysulfide species increases the C—S bond as well, for the same reason: charge delocalization along a sulfur chain.
5. In the polysulfide species PA, GPB, BDE, and HA all decrease with the increasing sulfur chain. The lower PA and GPB values indicate that in a biological system, polysulfide chains are more likely to exist as deprotonated rather than neutral species compared to the parent thiol or selenol.
6. To perform reliable thermochemical calculations, the  $\omega$ B97XD/6-311G(2d,p) level of theory was found to give the most accurate results relative to the benchmark and this was then used to evaluate the geometry and thermochemistry of the more complex cysteine and cysteine per/polysulfide species.

### 3.5 References

1. Zhong, L. W.; Holmgren, A. Essential Role of Selenium in the Catalytic Activities of Mammalian Thioredoxin Reductase Revealed by Characterization of Recombinant Enzymes with selenocysteine Mutations. *J. Biol. Chem.* **2000**, *275*, 18121–18128.
2. Fernandez, F. J.; Arda, A.; Lopez-Esteva, M.; Aranda, J.; Pena-Soler, E.; Garces, F.; Round, A.; Campos-Olivas, R.; Bruix, M.; Coll, M.; *et al.* Mechanism of Sulfur Transfer Across Protein-Protein Interfaces: The Cysteine Desulfurase Model System. *ACS Catal.* **2016**, *6*, 3975–3984.
3. Kahya, M. C.; Naziroglu, M.; Ovey, I. S. Modulation of Diabetes-Induced Oxidative Stress, Apoptosis, and  $\text{Ca}^{2+}$  Entry Through TRPM2 and TRPV1 Channels in Dorsal Root Ganglion and Hippocampus of Diabetic Rats by Melatonin and Selenium. *Mol. Neurobiol.* **2017**, *54*, 2345–2360.
4. Paulsen, C. E.; Carroll, K. S. Cysteine-Mediated Redox Signaling: Chemistry, Biology, and Tools for Discovery. *Chem. Rev.* **2013**, *113*, 4633–4679.
5. Garcia-Santamarina, S.; Boronat, S.; Hidalgo, E. Reversible Cysteine Oxidation in Hydrogen Peroxide Sensing and Signal Transduction. *Biochemistry* **2014**, *53*, 2560–2580.
6. Huang, Z.; Rose, A. H.; Hoffmann, P. R. The Role of Selenium in Inflammation and Immunity: From Molecular Mechanisms to Therapeutic Opportunities. *Antioxid. Redox. Signal.* **2012**, *16*, 705–743.
7. Ida, T.; Sawa, T.; Ihara, H.; Tsuchiya, Y.; Watanabe, Y.; Kumagai, Y.; Suematsu, M.; Motohashi, H.; Fujii, S.; Matsunaga, T.; *et al.* Reactive Cysteine Persulfides and S-Polythiolation Regulate Oxidative Stress and Redox Signaling. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 7606–7611.
8. Kim, H. J.; Ha, S.; Lee, H.Y.; Lee, K.J. ROSics: Chemistry and Proteomics of Cysteine Modifications in Redox Biology. *Mass Spectrom. Rev.* **2015**, *34*, 184–208.
9. Jacobs, E. T.; Jiang, R.; Alberts, D. S.; Greenberg, E.R.; Gunter, E.W.; Karagas, M. R.; Lanza, E.; Ratnasinghe, L.; Reid, M. E.; Schatzkin, A.; *et al.* Selenium and

- Colorectal Adenoma: Results of a Pooled Analysis. *J. Natl. Cancer Inst.* **2004**, *96*, 1669–1675.
10. Wu, G.; Fang, Y. Z.; Yang, S.; Lupton, J. R.; Turner, N. D. Glutathione Metabolism and its Implications for Health. *J. Nutr.* **2004**, *134*, 489–492.
  11. Giles, G. I.; Tasker, K. M.; Jacob, C. Hypothesis: The Role of Reactive Sulfur Species in Oxidative Stress. *Free Radic. Biol. Med.* **2001**, *31*, 1279–1283.
  12. Toohey, J. I. Sulfur Signaling: Is the Agent Sulfide or Sulfane? *Anal. Biochem.* **2011**, *413*, 1–7.
  13. Kimura, H. Physiological Role of Hydrogen Sulfide and Polysulfide in the Central Nervous System. *Neurochem. Int.* **2013**, *63*, 492–497.
  14. Kaneko, Y.; Kimura, Y.; Kimura, H.; Niki, I. L-Cysteine Inhibits Insulin Release from the Pancreatic  $\beta$ -cell—Possible Involvement of Metabolic Production of Hydrogen Sulfide, a Novel Gasotransmitter. *Diabetes* **2006**, *55*, 1391–1397.
  15. Lo Faro, M. L.; Fox, B.; Whatmore, J. L.; Winyard, P. G.; Whiteman, M. Hydrogen Sulfide and Nitric Oxide Interactions in Inflammation. *Nitric Oxide Biol. Chem.* **2014**, *41*, 38–47.
  16. Yu, B.; Zheng, Y.; Yuan, Z.; Li, S.; Zhu, H.; De La Cruz, L. K.; Zhang, J.; Ji, K.; Wang, S.; Wang, B. Toward Direct Protein S-Persulfidation: A Prodrug Approach That Directly Delivers Hydrogen Persulfide. *J. Am. Chem. Soc.* **2018**, *140*, 30–33.
  17. Giles, G. I.; Nasim, M. J.; Ali, W.; Jacob, C. The Reactive Sulfur Species Concept: 15 Years On. *Antioxidants* **2017**, *6*, 38.
  18. Shen, X. G.; Peter, E. A.; Bir, S.; Wang, R.; Kevil, C. G. Analytical Measurement of Discrete Hydrogen Sulfide Pools in Biological Specimens. *Free Radic. Biol. Med.* **2012**, *52*, 2276–2283.
  19. Iciek, M.; Wlodek, L. Biosynthesis and Biological Properties of Compounds Containing Highly Reactive, Reduced Sulfane Sulfur. *Polish J. Pharmacol.* **2001**, *53*, 215–225.

20. Mishanina, A. V.; Libiad, M.; Banerjee, R. Biogenesis of Reactive Sulfur Species for Signaling by Hydrogen Sulfide Oxidation Pathways. *Nat. Chem. Biol.* **2015**, *11*, 457–464.
21. Akaike, T.; Ida, T.; Wei, F. Y.; Nishida, M.; Kumagai, Y.; Alam, M. M.; Ihara, H.; Sawa, T.; Matsunaga, T.; Kasamatsu, S.; *et al.* Cysteinyl-tRNA Synthetase Governs Cysteine Polysulfidation and Mitochondrial Bioenergetics. *Nat. Commun.* **2017**, *8*, 1177.
22. Koike, S.; Nishimoto, S.; Ogasawara, Y. Cysteine Persulfides and Polysulfides Produced by Exchange Reactions with H<sub>2</sub>S Protect SH-SY5Y Cells from Methylglyoxal-Induced Toxicity through Nrf2 Activation. *Redox Biol.* **2017**, *12*, 530–539.
23. Munchberg, U.; Anwar, A.; Mecklenburg, S.; Jacob, C. Polysulfides as Biologically Active Ingredients of Garlic. *Org. Biomol. Chem.* **2007**, *5*, 1505–1518.
24. Gruhlke, M. C. H.; Nicco, C.; Batteux, F.; Slusarenko, A. J. The Effects of Allicin, a Reactive Sulfur Species from Garlic, on a Selection of Mammalian Cell Lines. *Antioxidants* **2017**, *6*, 1.
25. Millikin, R.; Bianco, C. L.; White, C.; Saund, S. S.; Henriquez, S.; Sosa, V.; Akaike, T.; Kumagai, Y.; Soeda, S.; Toscano, J. P.; *et al.* The Chemical Biology of Protein Hydropersulfides: Studies of a Possible Protective Function of Biological Hydropersulfide Generation. *Free Radic. Biol. Med.* **2016**, *97*, 136–147.
26. Cerqueira, N. M.; Fernandes, P. A.; Gonzalez, P. J.; Moura, J. J.; Ramos, M. J. The Sulfur Shift: An Activation Mechanism for Periplasmic Nitrate Reductase and Formate Dehydrogenase. *Inorg. Chem.* **2013**, *52*, 10766–10772.
27. Cheng, Q.; Sandalova, T.; Lindqvist, Y.; Arner, E. S. Crystal Structure and Catalysis of the Selenoprotein Thioredoxin Reductase 1. *J. Biol. Chem.* **2009**, *284*, 3998–4008.
28. Searle, P. A.; Molinski, T. F. 5 New Alkaloids from the Tropical Ascidian, *Lissoclinum* Sp—Lissoclinotoxin-a Is Chiral. *J. Org. Chem.* **1994**, *59*, 6600–6605.

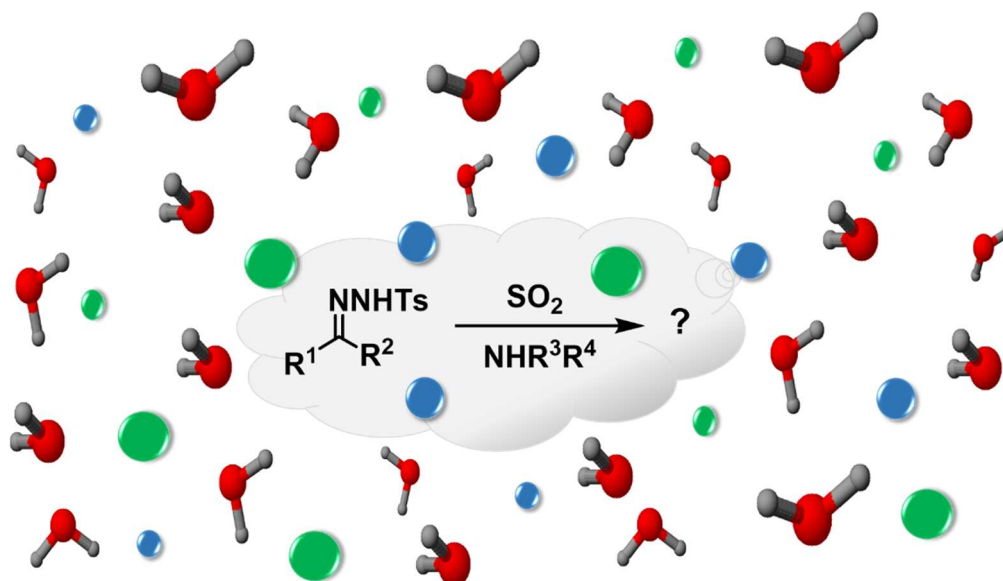
29. Brzostowska, E. M.; Greer, A. The Role of Amine in the Mechanism of Pentathiepin (Polysulfur) Antitumor Agents. *J. Am. Chem. Soc.* **2003**, *125*, 396–404.
30. Chauvin, J. P. R.; Griesser, M.; Pratt, D. A. Hydropersulfides: H-Atom Transfer Agents Par Excellence. *J. Am. Chem. Soc.* **2017**, *139*, 6484–6493.
31. Bortoli, M.; Torsello, M.; Bickelhaupt, F. M.; Orian, L. Role of the Chalcogen (S, Se, Te) in the Oxidation Mechanism of the Glutathione Peroxidase Active Site. *Chem. Phys. Chem.* **2017**, *18*, 2990–2998.
32. Bachrach, S. M.; Demoin, D. W.; Luk, M.; Miller, J.V. Nucleophilic Attack at Selenium in Diselenides and Selenosulfides. A Computational Study. *J. Phys. Chem. A* **2004**, *108*, 4040–4046.
33. Huang, G. T.; Yu, J. S. K. Enzyme Catalysis that Paves the Way for S-Sulphydration via Sulfur Atom Transfer. *J. Phys. Chem. B* **2016**, *120*, 4608–4615.
34. Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*, 2nd ed.; Wiley-VCH: Weinheim; New York, NY, USA, 2001.
35. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; N. Rega; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; J. V. Ortiz; Cioslowski, J.; Fox, D. J. *Gaussian 09*, E.01; Gaussian, Inc.: Wallingford CT, 2013.

36. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; J. Gao, N. R.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; S. S. Iyengar; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16*, Gaussian Inc.: Wallingford, CT, 2016.
37. Becke, A. D. A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.
38. Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. *Phys. Rev. B* **1988**, *37*, 785–789.
39. Perdew, J. P.; Wang, Y. Accurate and Simple Analytic Representation of the Electron-Gas Correlation-Energy. *Phys. Rev. B* **1992**, *45*, 13244–13249.
40. Jursic, B. S. Density Functional Theory and Complete Basis Set *Ab Initio* Evaluation of Proton Affinity for Some Selected Chemical Systems. *J. Mol. Struct. Theochem* **1999**, *487*, 193–203.
41. Bras, N. F.; Perez, M. A. S.; Fernandes, P. A.; Silva, P. J.; Ramos, M. J. Accuracy of Density Functionals in the Prediction of Electronic Proton Affinities of Amino Acid Side Chains. *J. Chem. Theory Comput.* **2011**, *7*, 3898–3908.

42. Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of four M06-Class Functionals and 12 other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
43. Zhao, Y.; Truhlar, D. G. Exploring the Limit of Accuracy of the Global Hybrid Meta Density Functional for Main-Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2008**, *4*, 1849–1868.
44. Chai, J. D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
45. Hunter, E. P. L.; Lias, S. G. Evaluated Gas Phase Basicities and Proton Affinities of Molecules: An update. *J. Phys. Chem. Ref. Data* **1998**, *27*, 413–656.
46. Moser, A.; Range, K.; York, D. M. Accurate Proton Affinity and Gas-Phase Basicity Values for Molecules Important in Biocatalysis. *J. Phys. Chem. B* **2010**, *114*, 13911–13921

## CHAPTER 4

# Computational Investigations Into Implicit and Explicit Roles of Solvent In Sulfonamide Formation





### 4.1 Introduction

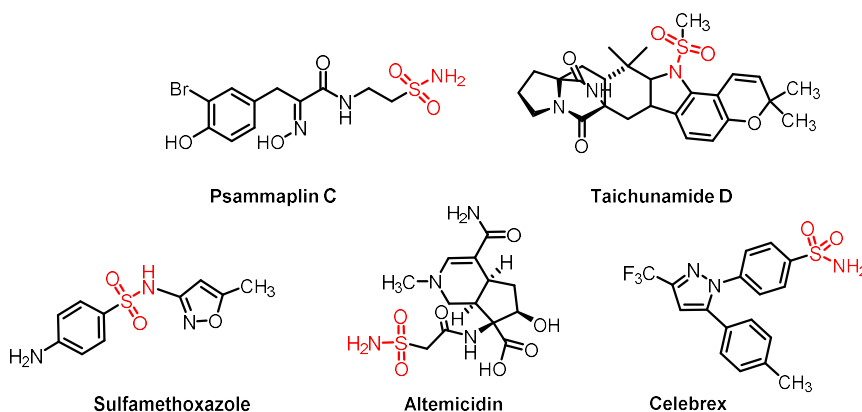
The environment, in particular solution, in which chemical reactants occur and encounter each other can play a central role in their reactivity and reaction. This is due to the one or more passive or active possible effects it may have on the molecules and their, for example, structure, physical and chemical properties including solubility and reactivity, as well as the actual mechanisms (including kinetics and thermodynamics) by which they may react.<sup>1</sup> That is, in solution, the solutes behavior is qualified by the solvent through their varied types of interactions.

There have been several studies on the impact and importance of solvent on a variety of chemical processes.<sup>2-9</sup> For example, Hawker *et al.* examined nucleophilic aromatic substitution reactions and concluded that increasing the presence of ionic liquid solvents enhanced the rate of reaction result due to stabilizing interactions between the solvents and reaction transition structures.<sup>8</sup> In addition, Notash *et al.* examined the reaction of cadmium iodide with selected polymers.<sup>9</sup> In particular, they showed that depending on the choice of solvent different one-dimensional coordination polymers, including a zigzag chain or sinusoidal wave, could be formed.<sup>9</sup> Meanwhile, Binh Nguyen *et al.* showed that the solvent choice can influence the stereo and regioselectivity of an acrylic acid derivative formed *via* photodimerization.<sup>10</sup>

Of increasing interest, from a broad range of perspectives, are molecular species that contain an S—N bond. Indeed, approximately 100 natural products are now known to contain such a bond or derivative thereof, *e.g.*, a sulfilimine bond (S=N).<sup>11</sup> Of the various possible chemical functionalities containing an S—N bond, however, those with a sulfonamide (RS(O)<sub>2</sub>-NR'R'') group have been of particular and significant interest. This is due in large part to their important roles in biology as well as their broad spectrum potential applications from pharmaceutical and medicinal chemistry to organic synthesis.<sup>12-14</sup> Indeed, they have been identified as a leading constituent of drug design and development due to their wide array of therapeutic and pharmacological activities (*e.g.*,

anti-inflammatory, anticancer, antioxidant, and antitumor), chemical and metabolic stability, and properties including polarity and hydrogen bonding capabilities, as well as their three-dimensional structures.<sup>15-21</sup> For instance, Psammaplin C, Taichunamide D, and Altemicidin are naturally occurring products (**Scheme 4.1**). Psammaplin C has shown potent inhibition of a cancer-associated carbonic anhydrase while Altemicidin has exhibited potential as an antitumor agent and herbicide. Meanwhile, Sulfamethoxazole and Celebrex are therapeutic agents used to treat bacterial infections and inflammation, respectively. Thus there is considerable interest in the medicinal and industrial potential of sulfonamide-containing natural products or synthetic agents.

**Scheme 4.1.** Schematic representative of selected natural and therapeutic sulfonamides species.<sup>11, 21</sup>



A variety of experimental synthetic pathways are presently known and used for the formation of such groups including insertion of SO<sub>2</sub> or its derivatives *via* nucleophilic substitution or radical processes to form S—N bonds and the RS(O)<sub>2</sub>—NR'R'' functionality.<sup>22-29</sup> Recently, Tsai *et al.* described for the first time a novel route for sulfonamide formation *via* the reaction of N-tosyl hydrazone, SO<sub>2</sub>, and amines under metal-free one-pot reaction conditions.<sup>30</sup> It was noted that varying the solvent resulted in different product yields. Furthermore, three possible mechanisms by which the product

might be formed were proposed. Unfortunately, however, how each possible pathway and its thermochemistry may be influenced by the choice of solvent remained unclear.

Computationally, a number of studies have examined the effects of solvent, either by implicit and/or explicit consideration, on the properties, chemistry, and reactions for a variety of chemical systems.<sup>31-40</sup> For example, recently Liu and co-workers investigated the reaction of the fluoride anion with ethyl bromide in an ethanol solution.<sup>39</sup> Notably, their study in part examined the effect of the solvent on the competition between  $S_N2$  substitutions and E2 elimination reactions. While in the gas-phase the E2 pathway was preferred, in solution solute-solvent interactions resulted in a preference for the  $S_N2$  nucleophilic displacement mechanism due in part to stabilization of its transition structure. In addition, recently the general effects of the polarity of the environment on the double proton transfers (DPTs) within the adenine-thymine and guanine-cytosine nucleobase pairs were investigated using DFT-based methods.<sup>40</sup> It was shown that the polarity of the environment itself could significantly impact the feasibility and mechanism of such processes. For instance, in solution both base pairs showed a preference for stepwise-DPTs while in the gas-phase thermodynamically feasible DPT was only observed for the guanine-cytosine base pair and *via* a concerted-DPT mechanism.

Developing simple, convenient and efficient synthetic strategies, such as the promising green-chemistry approach of one-pot syntheses, requires both an understanding of the reactants and target chemical systems, and also the role and effects of the environment and its constituents in which they occur. Hence, in the present study, we have applied DFT-based methods, in combination with a systematic series of chemical models, to investigate the potential effects and roles of the general reaction environment as well as that of individual solvent molecules on the mechanisms of formation of sulfonamide-containing compounds. To help facilitate comparison with experimental data, we have examined their formation of sulfonamide *via* reactions involving N-tosyl hydrazone,  $SO_2$ , and piperidine

as previously described by Tsai *et al.*<sup>30</sup> In particular, we have examined implicit and explicit effects of solvent and its constituent molecules on the possible mechanisms.

### 4.2 Computational Methods

All calculations were performed using the Gaussian 09 suite of programs.<sup>41</sup> Optimized geometries optimizations were obtained using the *meta*-GGA density functional theory (DFT) method M06-2X in combination with the 6-31G(d,p) basis set.<sup>42</sup> Except for gas-phase calculations, the polarity of the surrounding solvent environment was included using the IEFPCM solvation model. A wide range of solvent dielectric constants ( $\epsilon$ ), as provided in Gaussian 09, were used: 2.37 (toluene), 10.12 (dichloroethane), 35.68 (acetonitrile), 46.82 (DMSO), and 78.35 (water). It is noted that for gas-phase calculations  $\epsilon=1$ . Harmonic vibrational frequencies for all optimized geometries were obtained at the same levels of theory to ensure structures were indeed optimized and to characterize each as either an energy minimum or transition structure. To obtain more reliable and accurate relative energies, single-point calculations were performed at the M06-2X/6-311+G(2df,p) level of theory on the above optimized structures. As for optimizations and frequency calculations, for those single-points which included the polarity of the solvent the IEFPCM method was also used. Topological analyses of the electron densities of various complexes and species were performed using the AIMALL program.<sup>43</sup> It is noted that we chose the N-tosyl hydrazone/SO<sub>2</sub>/piperidine system as our model to examine sulfonamide formation as in the experimental study of Tsai *et al.*<sup>30</sup> as it provided the highest yield.

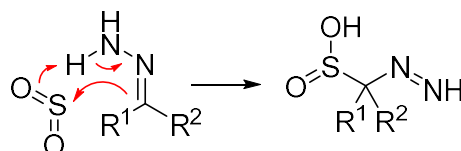
### 4.3 Results and Discussion

According to what has been proposed by Tsai *et al.*<sup>30</sup>, we have started studying the mechanism of sulfonamide formation through the insertion of SO<sub>2</sub> into N-tosyl hydrazone or the diaza compound generated *in situ* from the amine-initiated decomposition of the N-tosyl hydrazone.

### 4.3.1 Reactivity of SO<sub>2</sub> toward N-tosyl Hydrazone

Previously the direct ene reaction (also known as the Alder-ene reaction) of sulfur dioxide with alkenes to form  $\beta,\gamma$ -unsaturated sulfinic acid have been examined.<sup>44-46</sup> In the current study, we have also investigated the reactivity of SO<sub>2</sub> toward hydrazone as a potential enophile and ene compounds, respectively under the ene reaction (**Scheme 4.2**).

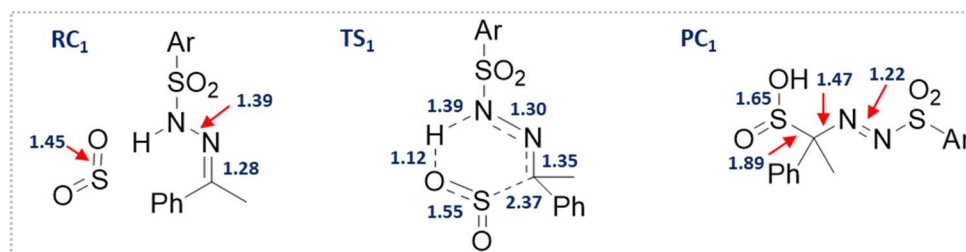
**Scheme 4.2.** The ene reaction of SO<sub>2</sub> and hydrazones.



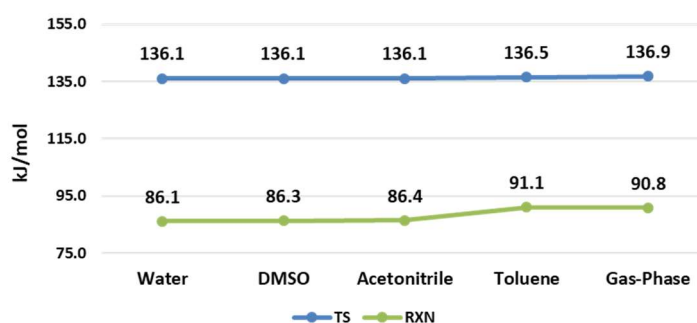
#### 4.3.1.1 Implicit Solvent Model

To elucidate the mechanism for the ene reaction of SO<sub>2</sub> with N-tosyl hydrazone, we have investigated the possible reaction pathway for the formation of  $\beta,\gamma$ -unsaturated sulfinic acid. Schematic illustration of the obtained optimized structures of the reactant complex (**RC<sub>1</sub>**), transition structure (**TS<sub>1</sub>**), and product complex (**PC<sub>1</sub>**) in the medium with a dielectric constant of 46.82 (DMSO) using the IEFPCM solvation model are shown in **Scheme 4.3**. As it is obvious, this reaction proceeds through a concerted pathway with the simultaneous double-bond migration from C=N to N=N and 1,5-H shift from N of N-tosyl hydrazone to O atom of SO<sub>2</sub>. The obtained 6-membered cyclic transition structure with the activation barrier of 136.1 kJ/mol leads to the formation of a  $\beta,\gamma$ -unsaturated sulfinic acid derivative (**PC<sub>1</sub>**) which lies 86.3 kJ/mol higher in energy relative to **RC<sub>1</sub>**.

**Scheme 4.3.** Schematic illustration of the optimized structures with the selected bond lengths in ångstroms (Å) for the reaction of SO<sub>2</sub> and N-tosyl hydrazone in DMSO.



Considering the influence of the environment on the obtained optimized structures and potential energy surface, different solvents with a quite varied dielectric constant of 78.35 (water), 35.68 (acetonitrile) and 2.37 (toluene) have been also applied implicitly. As **Figure 4.1** indicates, changing the solvent has no significant effects on the activation barrier and reaction energy. The only exception was observed using toluene which raises the reaction energy by  $\sim 5$  kJ/mol relative to other solvents with the higher polarity which is in agreement with the obtained reaction energy in the gas phase.



**Figure 4.1.** Calculated (electronic energy) activation barriers (difference in energy between reactants and transition structure; blue line) and reaction energies (difference in energy between reactants and products; green line) for the reaction of  $\text{SO}_2$  with N-tosyl hydrazone in various solvents and in the gas-phase.

For the electronic energies shown in **Figure 4.1**, we have also calculated the corresponding corrected Gibb's Free Energy values. The latter values are shown in **Appendix B4-1**. Both the electronic and corrected Gibb's Free Energies of reactions show the same overall trends. In contrast, unlike the essentially unchanging electronic energy activation barriers shown in **Figure 4.1**, the calculated Gibb's Free Energy activation barriers systematically decrease as the dielectric constant of the solvent decreases from that of water to the gas-phase. However, this decrease, 4.3 kJ/mol in total, is still relatively minor.

It is worth mentioning that the geometries of **RC**<sub>1</sub>, **TS**<sub>1</sub>, and **PC**<sub>1</sub> were not considerably different in water, DMSO and acetonitrile, while in toluene and in the gas-phase the newly formed C—S bond length has increased by ~0.06 to 1.95 Å compared to the changes in other solvents. This bond lengthening is consistent with the observation of H-bonding between H<sub>SOOH</sub> and O<sub>Ts</sub> which is not characterized in β,γ-unsaturated sulfinic acid by employing the other solvents with higher dielectric constant than toluene.

To investigate a more real physical picture of the reaction studied and the role of solvent molecules and its plausible intermolecular interactions with the reactant and product complexes and transition structures, a molecule of the different solvents has been explicitly included in the system. Concomitantly, the solvent environment has been implicitly modeled to take into account the general polarity. In the following discussion, the solvent molecule that has been explicitly included (*e.g.*, DMSO) is referred to as the base.

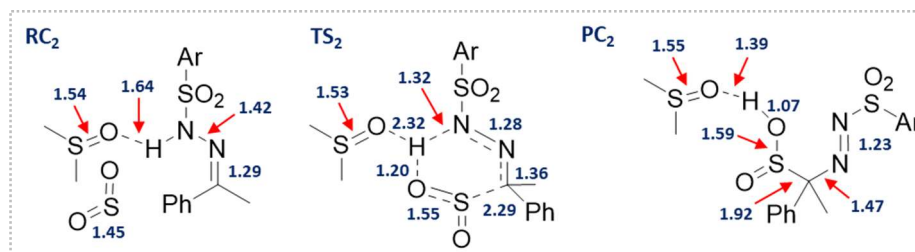
#### 4.3.1.2 Hybrid Explicit and Implicit Solvent Model

**DMSO-Based System:** **Scheme 4.4** represents the schematic optimized structures of **RC**<sub>2</sub>, **TS**<sub>2</sub>, and **PC**<sub>2</sub> of β,γ-unsaturated sulfinic acid formation in which one molecule of DMSO has been included in the system of interest. Obviously, the presence of explicit DMSO causes an H-bond formation between O atom of DMSO and H atom of N-tosyl hydrazone along increasing *r*(N—N) bond length by 0.03 to 1.42 Å compared to the implicit solvation model. DMSO also impacts on the transition structure by decreasing C···S bond distance from 2.37 to 2.29 Å. The QTAIM calculation values of electron density ( $\rho$ ) and Laplacian of the electron density ( $\nabla^2\rho$ ) at the bond critical point (BCP) for the C···S interaction are 0.072 and 0.031 a.u., respectively which compared to the corresponding values obtained for the **TS**<sub>1</sub> with calculated electron density ( $\rho$ ) of 0.060 and Laplacian of the electron density ( $\nabla^2\rho$ ) of 0.044 a.u. implies the stronger C···S closed-shell interaction. Also, the oxygen atom of DMSO forms an H-bond with the transferring proton from N to O atom with  $\rho$  and  $\nabla^2\rho$  of 0.014 and 0.049 a.u., respectively at the BCPs.

The obtained **TS<sub>2</sub>** leads to the formation of desired sulfinic acid (**PC<sub>2</sub>**) in which its  $r(\text{C—S})$  is lengthened by 0.03 to 1.92 Å relative to **PC<sub>1</sub>**. The calculated values of  $\rho$  and  $\nabla^2\rho$  at the BCP for C—S interaction are measured by 0.159 and  $-0.208$  a.u., respectively which is consistent with the formation of a weaker covalent interaction comparing previous solvation model with  $\rho$  and  $\nabla^2\rho$  of 0.171 and  $-0.258$  a.u., respectively (**PC<sub>1</sub>**). Nevertheless the strong H-bond of  $\text{H}_{\text{SOOH}}$  and  $\text{O}_{\text{DMSO}}$  with  $\rho$  and  $\nabla^2\rho$  of 0.101 and 0.083 contributes  $\beta,\gamma$ -unsaturated sulfinic acid's stability. It is worth mentioning  $r(\text{S—O})$  has been shortened by 0.06 to 1.59 Å relative the corresponding bond in the implicit solvation model (**PC<sub>1</sub>**).

All these types of interactions with DMSO result in the formation of  $\beta,\gamma$ -unsaturated sulfinic acid with 38.0 kJ/mol less energy than the previous solely implicit model. Although, **TS<sub>2</sub>** lies 5.2 higher in energy relative to the **TS<sub>1</sub>** without the presence of DMSO at 141.3 kJ/mol.

**Scheme 4.4.** Schematic illustration of the optimized structures with the selected bond lengths in ångstroms (Å) for the reaction of  $\text{SO}_2$  and N-tosyl hydrazone in the DMSO-based system.

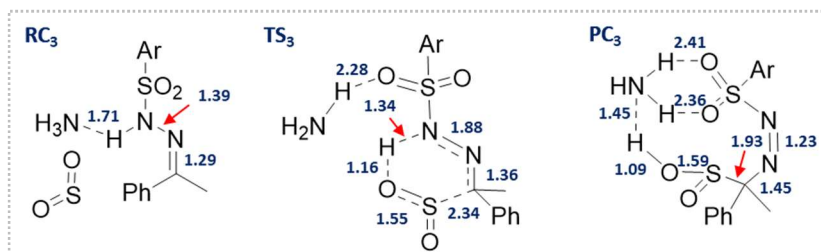


***NH<sub>3</sub> Based-System:*** Considering the effects of the base, a molecule of  $\text{NH}_3$  as a weak base has also been taken into account in the present study. The presence of  $\text{NH}_3$  in the **RC<sub>3</sub>** results in H-bond formation between the H atom of N-tosyl hydrazone and the N atom of ammonia with decreasing  $r(\text{N—N})$  bond length to 1.39 Å relative to **RC<sub>2</sub>** in DMSO-based system (see **Scheme 4.5**). In the obtained **TS<sub>3</sub>** two closed-shell interactions between  $\text{NH}_3\text{N}\cdots\text{N}_{\text{Hyrazon}}$  and  $\text{NH}_3\text{N}\cdots\text{S}_{\text{SO}_2}$  with measured  $\rho$  and  $\nabla^2\rho$  of 0.008 and 0.029, and 0.009 and 0.036 a.u., respectively at the BCP are observed. These values lie within the ranges for



$\rho$  (0.002-0.035 a.u.) and  $\nabla^2\rho$  (0.024-0.139 a.u.) previously suggested as a typical H-bonding.<sup>47</sup> In addition, an H-bonding with  $\rho$  and  $\nabla^2\rho$  of 0.014, 0.045 a.u., respectively at BCP for  $\text{NH}_3\text{H}\cdots\text{O}_{\text{Ts}}$  is characterized in the obtained **TS**<sub>3</sub>. Comparing with DMSO-based system,  $r(\text{C}\cdots\text{S})$  has been lengthened by 0.05 Å to 2.34, suggesting a weaker closed-shell interaction confirmed by its calculated  $\rho$  and  $\nabla^2\rho$  at its BCP by 0.064 and 0.040 a.u., respectively. Obtained **TS**<sub>3</sub> structure with an activation barrier of 133.9 kJ/mol, by ~2 kJ/mol less than the implicit solvation model and ~7 kJ/mol less than DMSO-based systems, results in  $\beta,\gamma$ -unsaturated sulfinic acid formation which lies 29.9 kJ/mol higher in energy relative to **RC**<sub>3</sub>. Indeed, formation of two strong H-bonds between  $\text{TsO}\cdots\text{H}_{\text{NH}_3}$  and  $\text{SOOH}\text{H}\cdots\text{N}_{\text{NH}_3}$  causes decreasing the energy of the reaction by 56.2 and 18.1 kJ/mol relative to the implicit solvation model and DMSO-based reaction. According to the QTAIM analysis no significant changes in the strength of the newly formed C—S was observed comparing to DMSO-based system.

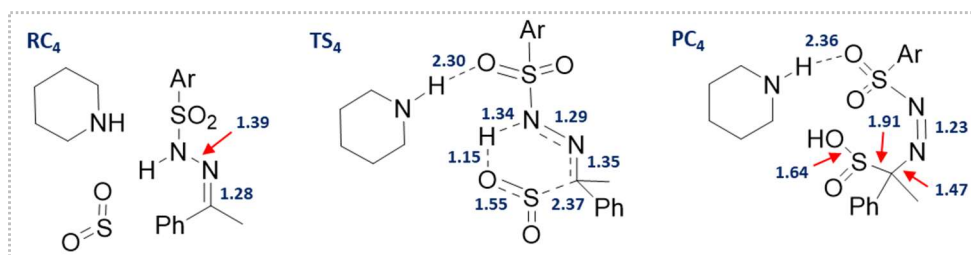
**Scheme 4.5.** Schematic illustration of the optimized structures with the selected bond lengths in ångstroms (Å) for the reaction of  $\text{SO}_2$  and N-tosyl hydrazone in the  $\text{HN}_3$ -based system.



**Piperidine-Based System:** The presence of piperidine as the base has also been investigated. Geometrical analysis indicates  $r(\text{C}\cdots\text{S})$  in the obtained **TS**<sub>4</sub> has been lengthened by 0.08 and 0.03 to 2.37 with the respect to the corresponding bond distance in DMSO- and  $\text{NH}_3$ -based systems (**Scheme 4.6**). The QTAIM results also confirm the weaker closed-shell interaction for  $\text{C}\cdots\text{S}$  in the **TS**<sub>5</sub>. Moreover, strong H-bonding between  $\text{N-H}_{\text{piperidine}}$  and  $\text{O}_{\text{Ts}}$ , some other weak closed-shell interactions including  $\text{TsO}\cdots\text{C}_\alpha/\text{H}_\alpha$ , and

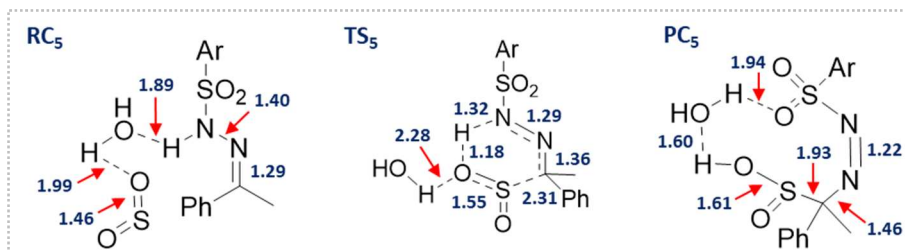
$\text{SO}_2\text{S}\cdots\text{H}_{\alpha/\beta}$  of piperidine are characterized in the obtained **TS<sub>4</sub>** structure. With a cost of 142.3 kJ/mol, transition structure yields the sulfinic acid derivative which stands 97.8 kJ/mol higher in energy relative to **RC<sub>4</sub>**. Despite formation as strong as C—S in different aforementioned systems, the presence of piperidine increases the energy of the reaction by 49.5, 67.9 and 11.5 kJ/mol relative to DMSO-,  $\text{NH}_3$ -based systems, and implicit solvation model, respectively.

**Scheme 4.6.** Schematic illustration of the optimized structures with the selected bond lengths in ångströms (Å) for the reaction of  $\text{SO}_2$  and N-tosyl hydrazone in the piperidine-based system.



**Water-Based System:** As **Scheme 4.7** shows presence of a molecule of water in the system of interest causes the formation of two strong hydrogen bonds with the H atom of N-tosyl hydrazone and O atom of sulfur dioxide. In the obtained **TS<sub>5</sub>** also a closed-shell interaction type between  $\text{wH}\cdots\text{O}_{\text{SO}_2}$  with  $\rho$  and  $\nabla^2\rho$  of 0.014 and 0.055 a.u., respectively at BCP has been characterized. Both QTAIM calculation and  $r(\text{C}\cdots\text{S})$  implies such a weaker interaction between C and S atom in the **TS<sub>5</sub>** compared to DMSO-based reaction. The obtained **TS<sub>5</sub>** with a cost of 149.7 kJ/mol leads to sulfinic acid formation. Although the activation barrier and energy of reaction has been increased by 8.4 and 17.8 kJ/mol compared to DMSO-based reaction, the strength of newly formed C—S has not changed noticeably relative to other systems.

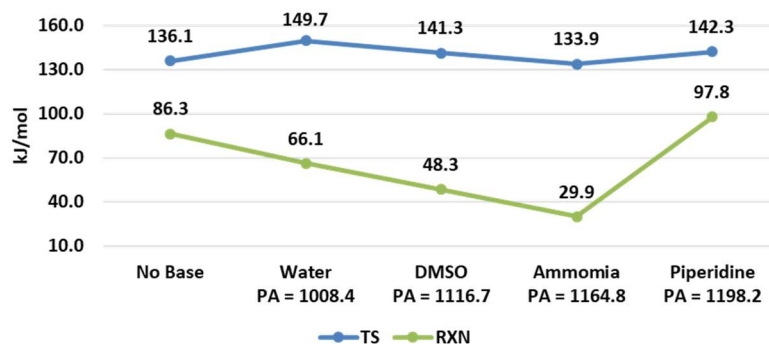
**Scheme 4.7.** Schematic illustration of the optimized structures with the selected bond lengths in ångströms (Å) for the reaction of SO<sub>2</sub> and N-tosyl hydrazone in the water-based system.



It should be mentioned that based on the calculated  $\rho$  and  $\nabla^2\rho$  at the BCPs resulted from QTAIM analysis, both transferring H and O atom of sulfur dioxide ( $\text{SO}_2\text{O}\cdots\text{H}$ ) and transferring H and N<sub>Hydrazone</sub> ( $\text{H}\cdots\text{N}_{\text{Hydrazone}}$ ) interactions of all the obtained **TS<sub>5</sub>** are characterized as the weak covalent bonds. Although  $\text{SO}_2\text{O}\cdots\text{H}$  interaction with measured  $\rho$  of 0.185 – 0.214 a.u. and  $\nabla^2\rho$  of –0.452 – –0.753 a.u. should be considered as the stronger covalent interaction compared to the  $\text{H}\cdots\text{N}_{\text{Hydrazone}}$  interaction.

**Figure 4.2** briefly represents the obtained activation barriers and reaction energies of the reactivity of SO<sub>2</sub> toward N-tosyl hydrazone in the DMSO with a dielectric constant of 46.82 in the presence of different based-systems. Obviously, presence of base raises the activation barrier compared the environment without any bases by 13.6 to 149.7 kJ/mol in the water-based reaction which has the lowest proton affinity in DMSO while by increasing the proton affinity of employed bases, the activation barrier decreases to 133.9 kJ/mol in the NH<sub>3</sub>-based reaction.

On the other hand, presence of base lowers the energy of the reaction from 86.3 in the environment without any base to 29.9 kJ/mol in NH<sub>3</sub>-based system. Indeed, decreasing the energy of the reaction is consistent with increasing the proton affinities of applied bases. The only exception was observed in the piperidine-based system with the highest proton affinity.

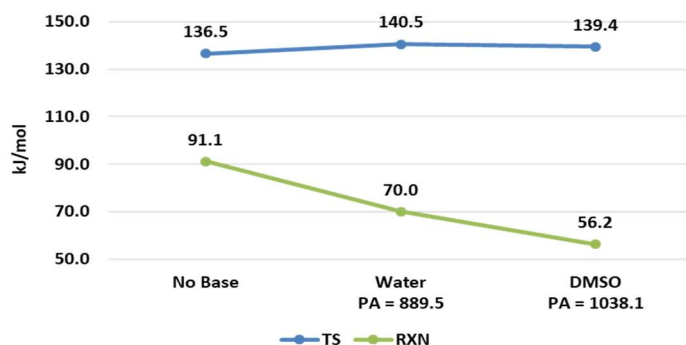


**Figure 4.2.** Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of  $\text{SO}_2$  toward N-tosyl hydrazone in the presence of different bases with the calculated proton affinity (PA) of each applied base (kJ/mol) in DMSO.

Expectedly, based on the obtained results from the implicit solvation model, negligible changes in geometrical parameters and calculated energies were observed by replacing the DMSO with water, as the implicit solvent with the higher dielectric constant of 78.35.

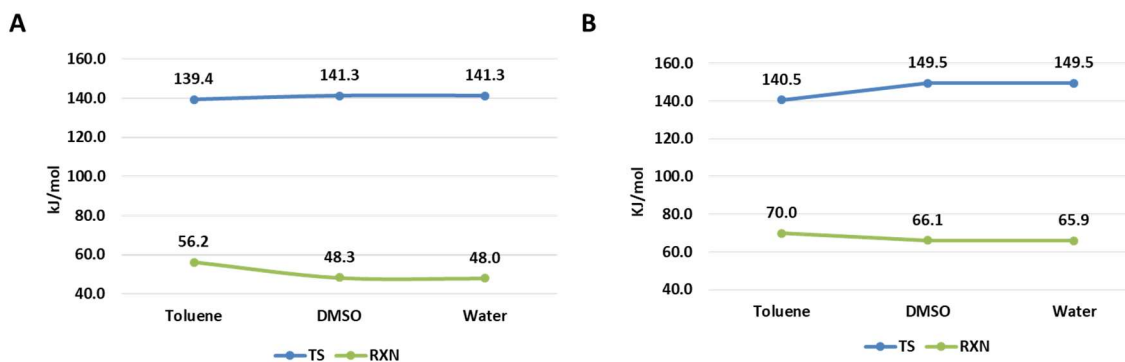
Effects of a non-polar environment by applying toluene with a dielectric constant of 2.37 as an implicit solvent, in the presence of water and DMSO as the base, are also investigated and shown in **Figure 4.3**. Clearly, a noticeable change is attributed to decreasing the energy of the reaction by 34.9 and 21.1 kJ/mol in DMSO- and water-based system respectively, relative to the environment without any bases. Additionally, toluene causes decreasing the activation barrier by 9.2 to 140.5 kJ/mol and 1.9 to 139.4 kJ/mol in water- and DMSO-based environment, respectively relative to the corresponding reaction in implicit DMSO in the presence of water and DMSO, respectively (see **Figure 4.2**). This is while toluene increases the energy of the reaction by 3.9 to 70.0 kJ/mol and 7.9 to 56.2 kJ/mol relative to the analogous reaction in the implicit DMSO model in the presence of water and DMSO, respectively (see **Figure 4.2**). As it is obvious, employing the base with higher proton affinity, has a considerable impact on the decreasing reaction energy with slight changes in the activation barrier. It is worth mentioning that the observed energy changes from water- to DMSO-based systems in toluene is less than energy changes in

DMSO as an implicit solvent; 1.1 vs. 8.4 kJ/mol for the activation barrier and 13.8 vs. 17.8 kJ/mol for the reaction energy.



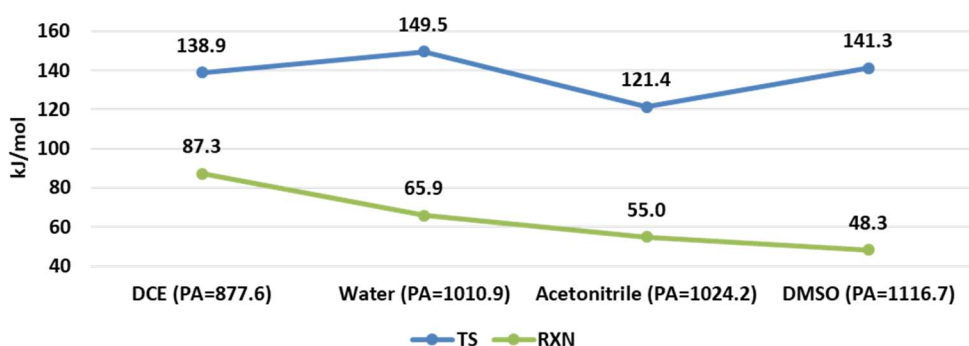
**Figure 4.3.** Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of  $\text{SO}_2$  toward N-tosyl hydrazone in the presence of water and DMSO as the base with the calculated proton affinity (PA) of each applied base (kJ/mol) in toluene.

To clarify the influence of the applied base, **Figure 4.4** indicates the energy changes by applying the constant bases including water and DMSO in an environment with a different dielectric constant from 2.37 (toluene) and 46.82 (DMSO), to 78.35 (water). Obviously, DMSO can be considered as a better base relative to water in proceeding the ene reaction between  $\text{SO}_2$  and N-tosyl hydrazone. It is noteworthy that changing the polarity of employed implicit solvent cannot be significantly taken into account especially from  $\epsilon=46.82$  (DMSO) to  $\epsilon=78.35$  (water).



**Figure 4.4.** Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of  $\text{SO}_2$  toward N-tosyl hydrazone in toluene, DMSO, and water with applying **A)** DMSO **B)** water as the base.

Alternatively, the influence of applying the same implicit and explicit solvent molecule in the reaction of  $\text{SO}_2$  and N-tosyl hydrazone is shown in **Figure 4.5**. Expectedly, by improving the proton affinity of the employed solvent the reaction energy has been remarkably decreased from 87.3 kJ/mol in DCE to 48.3 kJ/mol in DMSO. However, this decreasing is not consistent with the calculated activation barriers and the least one is obtained using the acetonitrile by 121.4 kJ/mol.

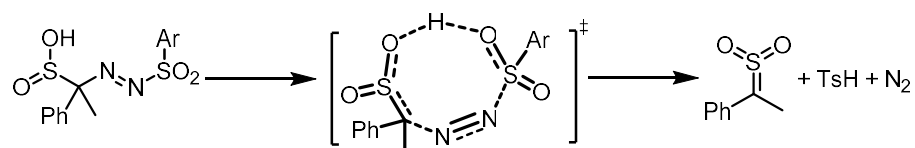


**Figure 4.5.** Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of  $\text{SO}_2$  toward N-tosyl hydrazone in an environment with the same implicit and explicit solvent molecule with the calculated proton affinity (PA) in kJ/mol.

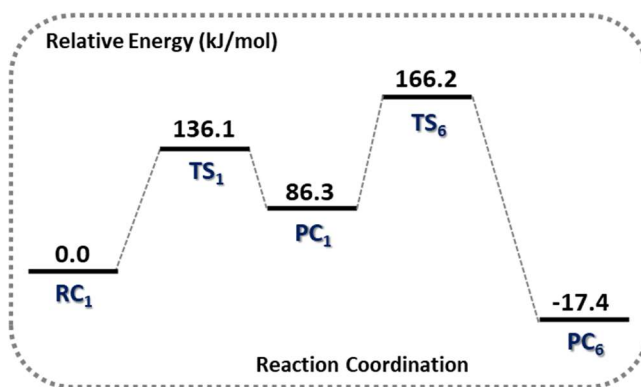
To form sulfene, which has the potential to undergo further reaction with amines to generate sulfonamide, decomposition of  $\beta,\gamma$ -unsaturated sulfinic acid was also studied. As

shown in **Scheme 4.8**, decomposition reaction proceeds *via* formation an 8-membered cyclic transition structure with a cost of 79.9 kJ/mol to produce sulfene, N<sub>2</sub> and TsH which lies 103.9 kJ/mol lower in energy relative to the sulfinic acid derivative. It is noteworthy that a water molecule contribution in the transition structure through the formation of a 10-membered cyclic transition structure results in increasing the activation barrier by 8.4 to 88.3 kJ/mol.

**Scheme 4.8.** Decomposition reaction of  $\beta,\gamma$ -unsaturated sulfinic acid to sulfene.



The obtained potential energy surface (PES) of  $\beta,\gamma$ -unsaturated sulfinic acid formation (**PC<sub>1</sub>**) from the reaction of SO<sub>2</sub> and N-tosyl hydrazone (**RC<sub>1</sub>**) and its decomposition to sulfene, TsH and N<sub>2</sub> (**PC<sub>6</sub>**) in DMSO is shown in **Figure 4.6**. Changing the polarity from 46.82 (DMSO) to 78.35 (water) has an insignificant influence on the obtained geometrical parameters and PES (data not shown).

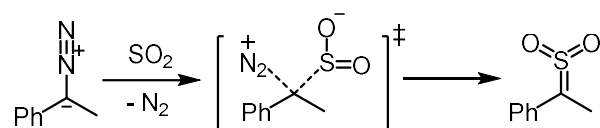


**Figure 4.6.** Potential energy surface for the formation of  $\beta,\gamma$ -unsaturated sulfinic acid and its decomposition to sulfene in DMSO.

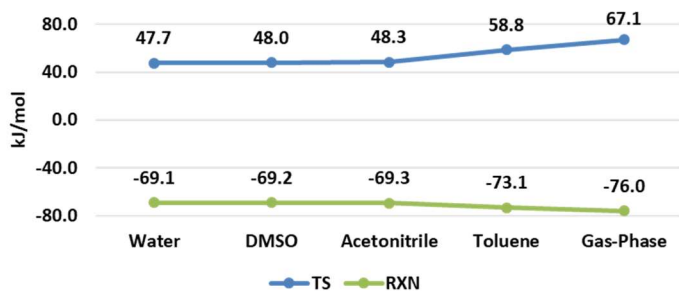
### 4.3.2 Reactivity of SO<sub>2</sub> toward Diaza Species

The diazo species generated *in situ* from the amine-initiated decomposition of the N-tosyl hydrazone can behave as a nucleophile in reaction with SO<sub>2</sub>. A one-step reaction of sulfene formation from the diazo species is shown in **Scheme 4.9**. The reaction proceeds through a simultaneous C⋯S and C⋯N bond formation and dissociation, respectively.

**Scheme 4.9.** Reactivity of diazo species toward SO<sub>2</sub> to form sulfene.



Solvents with a varied dielectric constant from 78.35 to 2.37 and then gas phase have been used to consider their impact on the geometrical parameters and energy values of the above reaction. As shown in **Figure 4.7**, employing solvents with a dielectric constant of 78.35 (water), 46.82 (DMSO), and 35.66 (acetonitrile), the observed changes in the energy values are negligible. Also, geometrical parameters represent insignificant changes using these solvents. Although a nonpolar solvent (toluene) and gas-phase increase the activation barrier by ~11 and 19 to 58.8 and 67.1 kJ/mol, respectively. The mentioned nonpolar environments cause lowering the energy of the reaction by < 7 kJ/mol relative to polar solvents.



**Figure 4.7.** Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of SO<sub>2</sub> toward diaza species in various solvents and in the gas-phase.

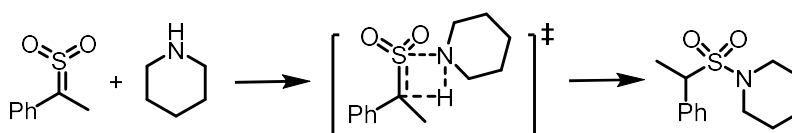


### 4.3.3 Reactivity of Sulfene toward Amines

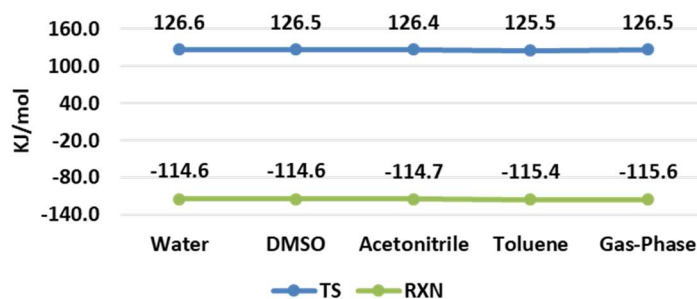
Sulfene resulted from the aforementioned reactions can undergo further reaction with amines to form sulfonamides. Piperidine as the amine with the highest reported yield by Tsai *et al.*<sup>30</sup> has been selected in this study for the formation of the desired product.

According to **Scheme 4.10**, solvent-unassisted addition reaction proceeds through a concerted one-step mechanism of S—N bond formation and proton transferring from amine (piperidine) to C atom of sulfene to form the sulfonamide.

**Scheme 4.10.** Reactivity of sulfene toward piperidine to form sulfonamide.



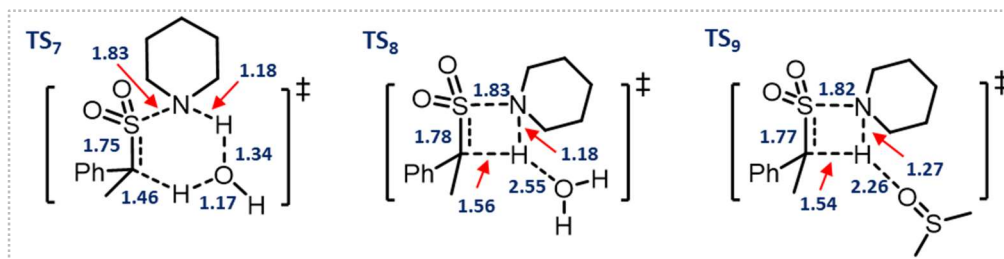
Formation a 4-membered cyclic transition structure with a cost of ~126 kJ/mol yields sulfonamide with the reaction energy of ~ -115 kJ/mol in an environment with different polarity from water to the gas phase (**Figure 4.8**).



**Figure 4.8.** Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of sulfene toward piperidine in various solvents and in the gas-phase

As expected and shown in **Scheme 4.11**, explicit solvent molecules potentially affect proton transferring from the N atom of piperidine to the C atom of sulfene actively *via* simultaneous accepting and donating proton and formation a 6-membered cyclic transition structure (**TS<sub>7</sub>**), or passively through H-bonding to the transferring proton (**TS<sub>8</sub>** and **TS<sub>9</sub>**).

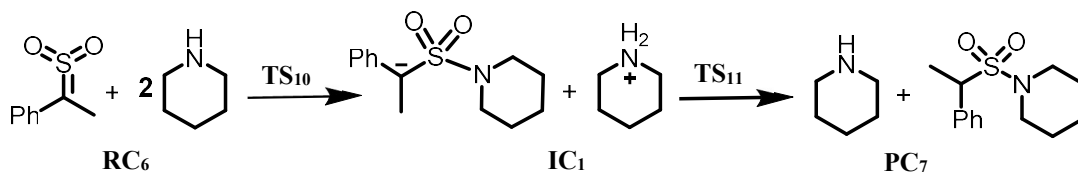
**Scheme 4.11.** Schematic description of the obtained transition structures with the selected bond lengths in ångstroms (Å) of solvent-assisted addition reaction of piperidine to sulfene in DMSO with the active role of the water molecule (**TS<sub>7</sub>**), the passive role of the water molecule (**TS<sub>8</sub>**), and the passive role of the DMSO (**TS<sub>9</sub>**).



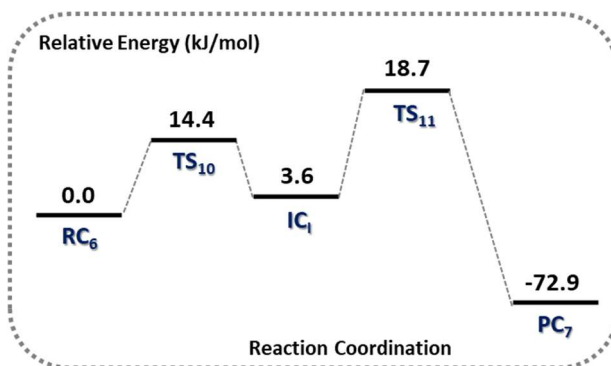
Based on the obtained potential energy surfaces, water molecule contributes to decreasing the activation barrier by 45.4 to 81.1 kJ/mol relative to the solvent-unassisted reaction by its active role in the formation of a 6-membered cyclic transition structure (**TS<sub>7</sub>**). However, due to the higher proton affinity of the N atom in piperidine compared to the O atom, water contribution through H-bonding with transferring proton causes increasing the activation barrier by 14.8 to 141.3 kJ/mol compared to the solvent-unassisted sulfonamide preparation (**TS<sub>8</sub>**). Also, the passive role of DMSO contribution toward sulfonamide formation has been studied. Similarly, presence of explicit DMSO molecule and its H-bonding with transferring proton increases the activation barrier by 10.1 to 136.6 kJ/mol results from higher proton affinity of piperidine relative to DMSO (**TS<sub>9</sub>**). Although DMSO with more proton affinity compared to water leads to the formation of a stronger H-bond with the transferring proton; 2.26 Å vs. 2.55 Å (see **Scheme 4.11**).

Alternatively, the effect of using two equivalents of piperidine acting as both base and nucleophile has also been investigated. As **Scheme 4.12** indicates the addition reaction undergoes through a 2-step mechanism to form sulfonamide.

**Scheme 4.12.** Formation of sulfonamide from the reaction of sulfene and two equivalents of piperidine.



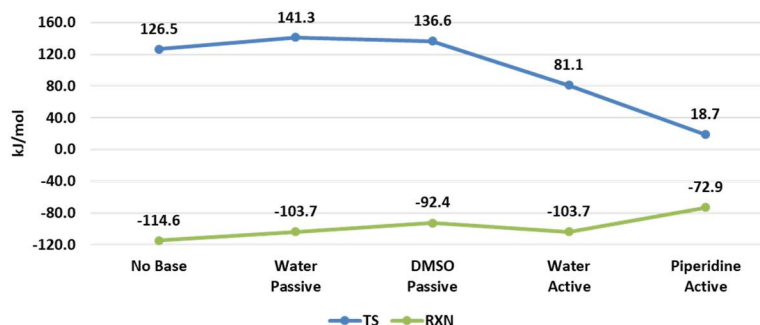
Due to higher proton affinity of piperidine relative to the water and DMSO, the first step of the reaction, coincident S—N bond formation and proton transferring to the second equivalent of piperidine, proceeds *via* 14.4 kJ/mol (**TS<sub>10</sub>**) to produce a deprotonated form of the desired sulfonamide (**IC<sub>1</sub>**). The second step of the reaction, as the rate-determining step, with a cost of 18.7 kJ/mol (**TS<sub>11</sub>**) results in sulfonamide which lies 72.9 kJ/mol lower in energy with respect to the **RC<sub>6</sub>** (see **Figure 4.9**). Performing the reaction in a more polar environment ( $\epsilon=78.35$ ) did not impact the structural parameters and PES considerably (data not shown).



**Figure 4.9.** Potential energy surface for the formation of sulfonamide from sulfene in the presence of two equivalents of piperidine in DMSO.

**Figure 4.10** summarizes the resulted energy values of the sulfonamide formation through the addition reaction of sulfene and piperidine with/without presence of a base. As described, applying bases passively with the lower proton affinity than piperidine cause increasing the activation barrier and energy of the reaction (*e.g.*, DMSO), while bases with their active roles (*e.g.*, water) are able to decrease the activation barrier noticeably. The

obtained results are strongly consistent with employing two equivalents of piperidine acting as both nucleophile and base that causes decreasing the activation barrier by  $\sim 7$  times relative to the result obtained applying implicit solvation model to form the desired sulfonamide.



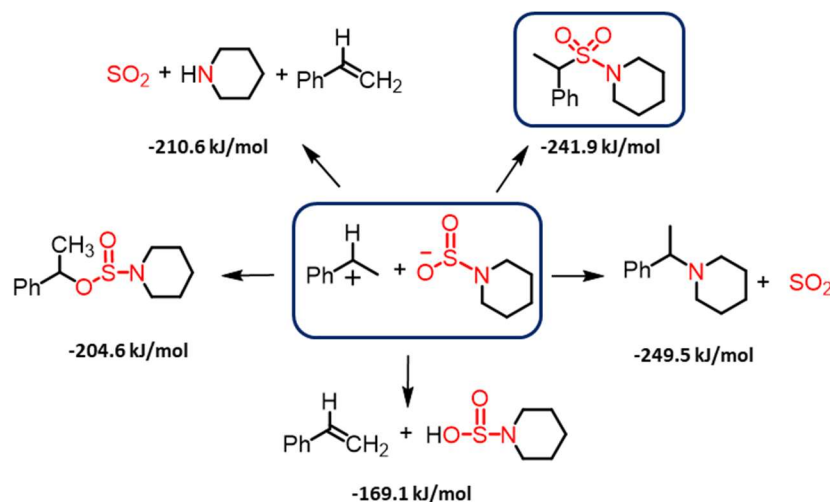
**Figure 4.10.** Calculated activation barriers (blue line) and reaction energies (green line) for the reactivity of sulfene toward piperidine with/without the presence of different bases in DMSO.

#### 4.3.4 Alternative Sulfonamide Formation from N-tosyl Hydrazone

The diaza intermediate from N-tosyl hydrazone decomposition can be potentially protonated in the reaction environment with the calculated  $\Delta E = -1133.0$  kJ/mol. This species is likely able to lose  $N_2$  with a low activation barrier of 3.0 kJ/mol to form  $CH_3CH^+Ph$  species. On the other hand, the nucleophilic (piperidine- $SO_2$ ) $^-$  complex, as a potential reactive moiety, is able to undergo further reaction with cationic  $CH_3CH^+Ph$  to form the desired sulfonamide ( $\Delta E = -241.9$  kJ/mol). Although some other products resulting from the nucleophilic potential of O, N, and S atoms with calculated NBO charges of  $-1.12$ ,  $-0.80$ , and  $1.64$  respectively, in (piperidine- $SO_2$ ) $^-$  complex or elimination reaction in which (piperidine- $SO_2$ ) $^-$  complex acts as the base are plausible (**Scheme 4.13**). As shown, two products from nucleophilic addition of S atom in (piperidine- $SO_2$ ) $^-$  complex to  $CH_3CH^+Ph$  ( $\Delta E = -241.9$  kJ/mol) and nucleophilic attack of the N atom in (piperidine- $SO_2$ ) $^-$  complex by losing  $SO_2$  to  $CH_3CH^+Ph$  ( $\Delta E = -249.5$  kJ/mol) are

considered as the most thermodynamically preferred products. It should be mentioned employing solvent with a higher dielectric constant ( $\epsilon=78.35$ ) increase the reaction energy by  $\sim 4$  kJ/mol with respect to the to DMSO as an implicit solvent.

**Scheme 4.13.** Plausible products of the reaction between (piperidine- $\text{SO}_2$ ) $^-$  complex and  $\text{CH}_3\text{CH}^+\text{Ph}$  in DMSO with corresponding formation energy.



#### 4.4 Conclusions

Under the impetus to investigate the significance of the reaction environment, solution-phase and individual solvent molecules and their atomistic impacts on the mechanism of sulfonamide formation, two computational solvation models, implicit and hybrid implicit/explicit solvation models, have been employed in this study.

Findings indicate that applying implicit solvents with different dielectric constants and various H-bond donor and acceptor capabilities has negligible effects on the transition structures obtained under the ene reaction between  $\text{SO}_2$  and N-tosyl hydrazine to form  $\beta,\gamma$ -unsaturated sulfinic acid. Although based on the QTAIM analysis the newly formed C—S bond is characterized as a bit stronger covalent bond in polar aprotic and protic solvents such as water, DMSO, and acetonitrile relative to the implicit nonpolar solvent such as toluene which is consistent with increasing the energy of the reaction.

On the other hand, results obtained from the hybrid implicit/explicit solvation model indicate that applying the base with a greater proton affinity, causes a noticeable decrease in energy of the ene reaction. The activation barrier with a slight decrease is less sensitive to the presence of the base. Furthermore in such based-systems, using the implicit nonpolar solvent *e.g.*, toluene leads to decreasing activation barrier compared to polar implicit solvents (DMSO and water), while it slightly increases the reaction energy. Comparing the DMSO- and water-based systems in different polar and nonpolar implicit solvents also highlights DMSO as a better base than water in the ene reaction.

Both decompositions of  $\beta,\gamma$ -unsaturated sulfinic acid and the reactivity of nucleophilic diazo species toward  $\text{SO}_2$  lead to sulfene formation which is almost independent of the effects of the applied polar protic and aprotic solvents. Although the results show that employing the nonpolar solvent increases the activation barrier of sulfene formation from the reaction of  $\text{SO}_2$  and diazo species consistent with a decrease in the reaction energy.

Energy values for the  $\text{S}(\text{O}_2)\text{—N}$  bond formation (desired sulfonamide) in the reaction of sulfene and amine (piperidine) are not significantly impacted by applying the implicit solvation model with different solvents and dielectric constants from 78.35 to 2.37.

In contrast, applying the hybrid implicit/explicit solvation model indicates that the presence of the base with its active contribution in proton transferring from the N atom of amine to the C atom of sulfene (*e.g.*, water) has a significant role in decreasing the activation barrier. However, considering the passive role of the applied bases with the lower proton affinity than amine causes an increase in the activation barrier compared to the corresponding value using the implicit solvation model. The obtained results are strongly consistent with employing two equivalents of piperidine acting as both nucleophile and base which decreases the activation barrier by  $\sim 7$  times relative to the result obtained applying implicit solvation model to form the desired sulfonamide.

#### 4.5 References

1. Reichardt, C., *Solvents and Solvent Effects in Organic Chemistry*. Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, 2010.
2. Ortega-Rojas, M.; Rivera-Ramírez, J.; Ávila-Ortiz, C.; Juaristi, E.; González-Muñoz, F.; Castillo, E.; Escalante, J., One-Pot Lipase-Catalyzed Enantioselective Synthesis of (R)-(-)-N-Benzyl-3-(benzylamino)butanamide: The Effect of Solvent Polarity on Enantioselectivity. *Molecules* **2017**, *22*, 2189.
3. Tirsoaga, A.; Cojocaru, B.; Teodorescu, C.; Vasiliu, F.; Grecu, M. N.; Ghica, D.; Parvulescu, V. I.; Garcia, H., C–N Cross-Coupling on Supported Copper Catalysts: The Effect of the Support, Oxidation State, Base and Solvent. *J. Catal.* **2016**, *341*, 205-220.
4. Ellebracht, N. C.; Jones, C. W., Amine functionalization of cellulose nanocrystals for acid–base organocatalysis: surface chemistry, Cross-Linking, and Solvent Effects. *Cellulose* **2018**, *25*, 6495-6512.
5. Yang, Q.; Dai, G.-L.; Yang, Y.-M.; Luo, Z.; Tang, Z.-Y., Solvent Effects: Syntheses of 3,3-Difluorooxindoles and 3-Fluorooxindoles from Hydrazonoindolin-2-one by Selectfluor. *J. Org. Chem.* **2018**, *83*, 6762-6768.
6. Chandramouli, N.; El-Behairy, M. F.; Lautrette, G.; Ferrand, Y.; Huc, I., Polar Solvent Effects on Tartaric Acid Binding by Aromatic Oligoamide Foldamer Capsules. *Org. Biomol. Chem.* **2016**, *14*, 2466-2472.
7. Hu, X.; Yang, Y.; Shi, S.; Shan, Z., Effects of Solvent on Inclusion Complexation of a Chiral Dipeptide toward Racemic BINOL. *Chirality* **2016**, *28*, 784-788.
8. Hawker, R. R.; Haines, R. S.; Harper, J. B., The Effect of Varying the Anion of an Ionic Liquid on the Solvent Effects on a Nucleophilic Aromatic Substitution Reaction. *Org. Biomol. Chem.* **2018**, *16*, 3453-3463.
9. Notash, B.; Rezaei Kheirkhah, B., The Effect of Solvent on One-Dimensional Cadmium Coordination Polymers. *New J. Chem.* **2018**, *42*, 15014-15021.

10. Binh Nguyen, T.; Anh Nguyen, L.; Corbin, M.; Retailleau, P.; Ermolenko, L.; Al-Mourabit, A., Toward the Synthesis of Sceptrin and Benzosceptrin: Solvent Effect in Stereo- and Regioselective [2+2] Photodimerization and Easy Access to the Fully Substituted Benzobutane. *Eur. J. Org. Chem.* **2018**, *2018*, 5861-5868.
11. Petkowski, J. J.; Bains, W.; Seager, S., Natural Products Containing a Nitrogen-Sulfur Bond. *J. Nat. Prod.* **2018**, *81*, 423-446.
12. Chen, Y. T., Recent Functionalizations of Primary Sulfonamides. *Synthesis-Stuttgart* **2016**, *48*, 2483-2522.
13. Dorn, J. M.; Alpern, M.; McNulty, C.; Volcheck, G. W., Sulfonamide Drug Allergy. *Curr. Allergy Asthm. R.* **2018**, *18*, 38.
14. Rakesh, K. P.; Wang, S. M.; Leng, J.; Ravindar, L.; Asiri, A. M.; Marwani, H. M.; Qin, H. L., Recent Development of Sulfonyl or Sulfonamide Hybrids as Potential Anticancer Agents: A Key Review. *Anti-Cancer Agents Med. Chem.* **2018**, *18*, 488-505.
15. Supuran, C. T.; Innocenti, A.; Mastrolorenzo, A.; Scozzafava, A., Antiviral sulfonamide derivatives. *Mini-Rev. Med. Chem.* **2004**, *4*, 189-200.
16. Greig, I. R.; Coste, E.; Ralston, S. H.; van 't Hof, R. J., Development of triarylsulfonamides as novel anti-inflammatory agents. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 816-820.
17. Capasso, C.; Supuran, C. T., Sulfa and Trimethoprim-Like Drugs-Antimetabolites Acting as Carbonic Anhydrase, Dihydropteroate Synthase and Dihydrofolate Reductase Inhibitors. *J. Enzyme Inhib. Med. Chem.* **2014**, *29*, 379-387.
18. Carta, F.; Mannelli, L. D.; Pinard, M.; Ghelardini, C.; Scozzafava, A.; McKenna, R.; Supuran, C. T., A Class of Sulfonamide Carbonic Anhydrase Inhibitors with Neuropathic Pain Modulating Effects. *Bioorg. Med. Chem.* **2015**, *23*, 1828-1840.
19. Anusha, S.; Sinha, A.; Rajeev, C. P. B.; Chu, T. T. T.; Mathai, J.; Huang, X. M.; Fuchs, J. E.; Shivananju, N.; Bender, A.; Preiser, P. R.; Rangappa, K. S.; Basappa; Chandramohanadas, R., Synthesis, Characterization and in Vitro Evaluation of Novel



Enantiomerically-Pure Sulphonamide Antimalarials. *Org. Biomol Chem.* **2015**, *13*, 10681-10690.

20. Supuran, C. T., Carbonic Anhydrase Inhibition and the Management of Hypoxic Tumors. *Metabolites* **2017**, *7*, 48.

21. Zhao, C.; Rakesh, K. P.; Ravidar, L.; Fang, W. Y.; Qin, H. L., Pharmaceutical and Medicinal Significance of Sulfur (S<sup>VI</sup>)-Containing Motifs for Drug Discovery: A Critical Review. *Eur. J. Med. Chem.* **2019**, *162*, 679-734.

22. Malet-Sanz, L.; Madrzak, J.; Ley, S. V.; Baxendale, I. R., Preparation of Arylsulfonyl Chlorides by Chlorosulfonylation of *in Situ* Generated Diazonium Salts Using a Continuous Flow Reactor. *Org. Biomol. Chem.* **2010**, *8*, 5324-5332.

23. DeBergh, J. R.; Niljianskul, N.; Buchwald, S. L., Synthesis of Aryl Sulfonamides via Palladium-Catalyzed Chlorosulfonylation of Arylboronic Acids. *J. Am. Chem. Soc.* **2013**, *135*, 10638-10641.

24. Shavnya, A.; Coffey, S. B.; Smith, A. C.; Mascitti, V., Palladium-Catalyzed Sulfonation of Aryl and Heteroaryl Halides: Direct Access to Sulfones and Sulfonamides. *Org. Lett.* **2013**, *15*, 6226-6229.

25. Johnson, M. W.; Bagley, S. W.; Mankad, N. P.; Bergman, R. G.; Mascitti, V.; Toste, F. D., Application of Fundamental Organometallic Chemistry to the Development of a Gold-Catalyzed Synthesis of Sulfinates Derivatives. *Angew. Chem. Int. Ed.* **2014**, *53*, 4404-4407.

26. Deeming, A. S.; Russell, C. J.; Willis, M. C., Combining Organometallic Reagents, the Sulfur Dioxide Surrogate DABSO, and Amines: A One-Pot Preparation of Sulfonamides, Amenable to Array Synthesis. *Angew. Chem. Int. Ed.* **2015**, *54*, 1168-1171.

27. Deeming, A. S.; Russell, C. J.; Willis, M. C., Palladium(II)-Catalyzed Synthesis of Sulfinates from Boronic Acids and DABSO: A Redox-Neutral, Phosphine-Free Transformation. *Angew. Chem. Int. Ed.* **2016**, *55*, 747-750.

28. Eid, N.; Karamé, I.; Andrioletti, B., Straightforward and Sustainable Synthesis of Sulfonamides in Water under Mild Conditions. *Eur. J. Org. Chem.* **2018**, *2018*, 5016-5022.
29. Hofman, K.; Liu, N. W.; Manolikakes, G., Radicals and Sulfur Dioxide: A Versatile Combination for the Construction of Sulfonyl-Containing Molecules. *Chem-Eur. J.* **2018**, *24*, 11852-11863.
30. Tsai, A. S.; Curto, J. M.; Rocke, B. N.; Dechert-Schmitt, A.-M. R.; Ingle, G. K.; Mascitti, V., One-Step Synthesis of Sulfonamides from N-Tosylhydrazones. *Org. Lett.* **2016**, *18*, 508-511.
31. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*. John Wiley & Sons Ltd.; New York: 2004.
32. Monard, G.; Rivail, J.-L., Solvent Effects in Quantum Chemistry. In *Handbook of Computational Chemistry*, Leszczynski, J.; Kaczmarek-Kedziera, A.; Puzyn, T.; G. Papadopoulos, M.; Reis, H.; K. Shukla, M., Eds. Springer International Publishing: Cham, 2017; pp 727-739.
33. De Vetta, M.; Menger, M. F. S. J.; Nogueira, J. J.; González, L., Solvent Effects on Electronically Excited States: QM/Continuum versus QM/Explicit Models. *J. Phys. Chem. B* **2018**, *122*, 2975-2984.
34. Zhang, J.; Zhang, H.; Wu, T.; Wang, Q.; van der Spoel, D., Comparison of Implicit and Explicit Solvent Models for the Calculation of Solvation Free Energy in Organic Solvents. *J. Chem. Theory Comput.* **2017**, *13*, 1034-1043.
35. De Wispelaere, K.; Wondergem, C. S.; Ensing, B.; Hemelsoet, K.; Meijer, E. J.; Weckhuysen, B. M.; Van Speybroeck, V.; Ruiz-Martínez, J., Insight into the Effect of Water on the Methanol-to-Olefins Conversion in H-SAPO-34 from Molecular Simulations and *in Situ* Microspectroscopy. *ACS Catal.* **2016**, *6*, 1991-2002.
36. Inaba, S., Theoretical Study of Decomposition of Methanediol in Aqueous Solution. *J. Chem. Phys. Chem. A* **2015**, *119*, 5816-5825.

37. Li, Y.; Hartke, B., Assessing Solvation Effects on Chemical Reactions with Globally Optimized Solvent Clusters. *ChemPhysChem* **2013**, *14*, 2678-2686.
38. Widmer, D. R.; Schwartz, B. J., Solvents Can Control Solute Molecular Identity. *Nat. Chem.* **2018**, *10*, 910-916.
39. Liu, X.; Zhang, J.; Yang, L.; Hase, W. L., How a Solvent Molecule Affects Competing Elimination and Substitution Dynamics. Insight into Mechanism Evolution with Increased Solvation. *J. Am. Chem. Soc.* **2018**, *140*, 10995-11005.
40. Romero, E. E.; Hernandez, F. E., Solvent Effect on the Intermolecular Proton Transfer of the Watson and Crick Guanine–Cytosine and Adenine–Thymine Base Pairs: a Polarizable Continuum Model Study. *Phys. Chem. Chem. Phys.* **2018**, *20*, 1198-1209.
41. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; N. Rega; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; J. V. Ortiz; Cioslowski, J.; Fox, D. J. *Gaussian 09*, E.01; Gaussian, Inc.: Wallingford CT, 2015.
42. Zhao, Y.; Truhlar, D. G., The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215-241.

43. AIMAll (Version 17.11.14); Todd A. Keith, TK Gristmill Software: Overland Park KS, USA, 2017.
44. Marković, D.; Volla, C. M. R.; Vogel, P.; Varela-Álvarez, A.; Sordo, J. A., BC13-Mediated Ene Reaction of Sulfur Dioxide and Unfunctionalized Alkenes. *Chem-Eur. J.* **2010**, *16*, 5969-5975.
45. Vogel, P.; Turks, M. r.; Bouchez, L.; Marković, D.; Varela-Álvarez, A.; Sordo, J. Á., New Organic Chemistry of Sulfur Dioxide. *Acc. Chem. Res.* **2007**, *40*, 931-942.
46. Bouchez, L. C.; Dubbaka, S. R.; Turks, M.; Vogel, P., Sulfur Dioxide Mediated One-Pot, Three- and Four-Component Syntheses of Polyfunctional Sulfonamides and Sulfonic Esters: Study of the Stereoselectivity of the Ene Reaction of Sulfur Dioxide. *J. Org. Chem.* **2004**, *69*, 6413-6418.
47. Liu, H.; Gauld, J. W., Protonation of Guanine Quartets and Quartet Stacks: Insights from DFT studies. *Phys. Chem. Chem. Phys.* **2009**, *11*, 278-287.

## CHAPTER 5

# Computational Insights Into the Non-Enzymatic XiaH Facilitated Radical-Based Formation of SO<sub>2</sub>-Containing Compounds



## 5.1 Introduction

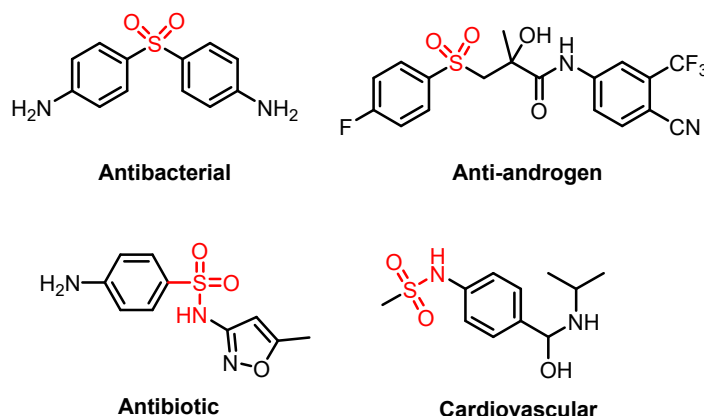
Within cells and organisms, sulfur dioxide (SO<sub>2</sub>) is endogenously biosynthesized *via* oxidation of either reduced sulfur compounds (*i.e.*, archaea and bacteria) or, for example, as in humans, sulfur-containing amino acids such as cysteine and homocysteine.<sup>1-2</sup> Importantly, SO<sub>2</sub> is now recognized as having diverse and important biological roles within cells and organisms. For example, in mammals<sup>3</sup> it has been shown to have antioxidant, anti-inflammatory, antihypertension activities as well as significant physiological roles in cardiac and blood vessel function and regulation.<sup>4</sup> Furthermore, in humans it is known as the fourth gasotransmitter.<sup>5</sup> However, conversely, it is also toxic in large quantities and has been implicated as having a role(s) in a number of pathological diseases, in particular, cardiovascular and respiratory diseases.<sup>6-7</sup> Indeed, as a result SO<sub>2</sub> has long been used as a preservative and antimicrobial agent in the food and beverage industries.<sup>8-9</sup>

Its diverse biological roles are due in part to its electronic structure and properties that allow it to act as both a nucleo- and electrophile, as well as also being able to also react *via* radical mechanisms.<sup>10</sup> These broad chemical capabilities have also been experimentally exploited in, for example, organic and polymer syntheses, and for medical applications such as therapeutic drugs.<sup>11-18</sup> Indeed, several leading pharmaceuticals with antibacterial, antiobesity, antitumor, antimalarial, and anti-neuropathic pain activities possess SO<sub>2</sub>-containing groups.<sup>16, 19</sup> This is due in large part to the fact that SO<sub>2</sub> can be reacted to form the versatile sulfone (R<sub>3</sub>C-SO<sub>2</sub>-CR<sub>3</sub>),<sup>20-23</sup> sulfonamide (R<sub>3</sub>C-SO<sub>2</sub>-NR<sub>2</sub>) and sulfamide (R<sub>2</sub>N-SO<sub>2</sub>-NR<sub>2</sub>) groups. These themselves can be functionalized with a diverse array of R groups ranging from simple hydrogen to bulkier and chemically diverse aliphatic, heterocyclic, or aromatic groups (**Scheme 5.1**). As such, they can possess significant biological properties and able to interact with, for example, proteins, metal ions or nucleic acids.<sup>18, 24-29</sup>

Despite their clear synthetic utility and potential, however, only a limited number of sulfonyl (-SO<sub>2</sub>-)-containing natural products have so far been identified.<sup>30</sup> Furthermore,

while there are several possible mechanisms by which sulfonyl (-SO<sub>2</sub>-) groups can be introduced into various compounds, there has been a long growing interest in radical-involving mechanisms.<sup>31-34</sup> Recently, Baunach *et al.*<sup>35</sup> observed the apparent enzyme-involved biosynthesis of bulky diaryl sulfonamide and diaryl sulfone antibiotics within the gram-positive *Streptomyces* bacteria of which several are known to produce antibiotics.<sup>36</sup> They proposed that these compounds are formed *via* a radical-based three-component reaction involving SO<sub>2</sub> and two equivalents of xiamycin (a pentacyclic indolosesquiterpene), and facilitated by the flavoenzyme *XiaH*.<sup>35</sup> More specifically, *XiaH* is thought to catalyze a single-electron oxidation of xiamycin (*Xia*) to generate a nitrogen-centered radical cation that then loses a proton to generate the corresponding neutral nitrogen-centered radical *Xia*•(N). It is noted that the actual mechanism of *XiaH* is itself still unclear.<sup>37</sup> Two carbon-centered radicals are also possibly formed on the adjacent benzyl rings *via* resonance stabilization. Subsequent reaction of these radicals with SO<sub>2</sub> yields N-SO<sub>2</sub> or C-SO<sub>2</sub> containing sulfonyl radical species which react further with additional xiamycin radicals to ultimately form the observed sulfone and sulfonamide products.

**Scheme 5.1.** Some examples of biologically active sulfone and sulfonamides.<sup>19</sup>



Computationally there have been a number of studies on related-systems that contain sulfur dioxide, as well as its reactions, and interactions especially with N-containing

compounds such as amines.<sup>38-45</sup> For example, Krupa *et al.* used the conventional correlated *ab initio* method MP2 and the dispersion-corrected hybrid-DFT method B3LYP-D3 to examine the interactions and properties of a series of HNCS $\cdots$ SO<sub>2</sub> complexes.<sup>44</sup> Their results indicated that van der Waals forces involving N $\cdots$ OS, N $\cdots$ SO<sub>2</sub>, S $\cdots$ SO<sub>2</sub> could be of greater importance for stabilizing some complexes than hydrogen bonds. In addition, a study by Sinha *et al.* examined in part, using the non-dispersion corrected B3LYP, the oxidation by SO<sub>2</sub> of *m*-xylene radicals generated *via* H-abstraction.<sup>45</sup> Notably, they concluded that SO<sub>2</sub> preferentially added to *m*-xylene radicals *via* one of its electronegative O-center's compared to its S-center. It is also noted that a number of computational studies have also examined the chemistry and properties of sulfone, sulfonamide and/or sulfamide compounds.<sup>46-49</sup> However, to date, few computational studies have examined mechanisms for formation of biomolecular sulfonyl radicals and/or SO<sub>2</sub>-crosslinked species involving C- or N-centered radicals.

In this present study we have applied conventional *ab initio* and density functional theory methods to study select reactions of SO<sub>2</sub> with xiamycin-related radicals. More specifically, we have examined reactions of SO<sub>2</sub> with a systematic series of xiamycin-inspired aromatic C- and N-centered radicals, in the gas-phase and aqueous solution, to form sulfonyl radicals or SO<sub>2</sub>-crosslinked sulfone, sulfonamide, and sulfamide species. This has also been done to provide insights into their structures as well as the effects of substituents and environment on the reactions. In addition, they also provide a greater understanding of the stability, reactivity, and nature of the  $\cdot\text{O}_2\text{S}-\text{C}$  or  $\cdot\text{O}_2\text{S}-\text{N}$  bonds formation.

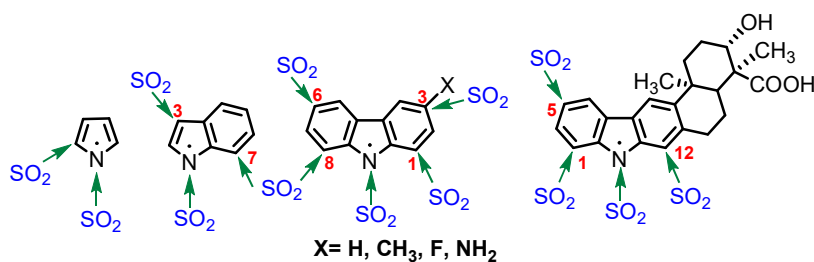
## 5.2 Computational Methods

All calculations were performed using the Gaussian 09<sup>50</sup> program. Optimized geometries optimizations were obtained using the conventional *ab initio* electron correlated MP2 method, the hybrid density functional theory (DFT) method B3LYP, composed of Becke's



three parameters (B3) exchange functional<sup>51</sup> in combination with the Lee, Yang, and Parr (LYP) correlation functional,<sup>52</sup> and the *meta*-GGA DFT method M06-2X,<sup>53</sup> in combination with the double-zeta and triple-zeta basis sets 6-31G(d) and 6-311G(d), respectively. The empirical D3 dispersion correction by Grimme<sup>54</sup> was also used to model the non-covalent and dispersion interactions more accurately combined with the B3LYP functional for a number of studied models. For all optimized structures the harmonic vibrational frequencies were calculated at the same level of theory to characterize and ensure they were all energy minima (all frequencies real-valued). To obtain more accurate and reliable relative energies, single-point energy calculations were performed on the above optimized structures, and using the above methods but in combination with the more extensive 6-311+G(2df,p) basis set. For calculations in which the effects of the polarity of an aqueous solution was also included, the IEFPCM solvation model with a dielectric constant ( $\epsilon$ ) of 78.35 was used. For several gas-phase complexes, the basis set superposition error (BSSE) using the counterpoise method was determined and their interaction energy corrected.<sup>55</sup> The systematic series of xiamycin-inspired radicals shown in **Scheme 5.2**, *i.e.*, from an N-centered pyrrole radical to an N-centered xiamycin radical, were used to examine the effect of substituents on the reactions and properties of the resulting sulfonyl-radicals species.

**Scheme 5.2.** Schematic initial illustration of the radical species considered in this study.



### 5.3 Results and Discussion

The reactions that were examined in this study were the addition of SO<sub>2</sub> to an N-centered (un)substituted radical structures as shown in **Scheme 5.2**.

### 5.3.1 Pyrrolyl

The highest calculated spin density on the pyrrolyl ring is related to the two C atoms adjacent to the N atom with the value of 0.53, while the spin density on the N atom is computed  $-0.17$ . In following, inserting the SO<sub>2</sub> molecule into the pyrrolyl ring through the formation of N—S(O<sub>2</sub>) and C—S(O<sub>2</sub>) are discussed (see **Scheme 5.2**).

#### 5.3.1.1 N—S(O<sub>2</sub>) Bond Formation

In the obtained sulfonyl radical species, a product of the reaction between the S atom in sulfur dioxide and N atom in pyrrolyl using B3LYP and M06-2X functionals, the highest spin density is related to the S atom with the value of 0.33 and 0.38 respectively, which comparing them with MP2 represents a better performance of M06-2X. The second and third highest calculated spin density in the obtained species after S atom are attributed to the O atom of SO<sub>2</sub> and N atom of the ring respectively, confirmed by B3LYP, M06-2X, and MP2 functionals.

In addition, S—N bond length has been examined using B3LYP and M06-2X/6-31G(d) in fully optimized structures which are measured by 1.735 and 1.700 Å, respectively. By comparison with MP2, a better performance of M06-2X functional (by 0.012 Å difference) than B3LYP (by 0.035 Å difference) in  $r(\text{S—N})$  is observed. To consider the effect of the polar environment, S—N bond length has also been examined applying IEFPCM solvation model with a dielectric constant ( $\epsilon$ ) of 78.35. Results represent that  $r(\text{S—N})$  has been shortened only by  $< 0.006$  Å using B3LYP, M06-2X, and MP2. On the other hand, improving the valence description from double to triple-zeta, (6-31G(d) to 6-311G(d)), does not impact  $r(\text{S—N})$  significantly. As well, the inclusion of dispersion interaction effects *via* the use of Grimme's correction (B3LYP-D3) and its effect on the S—N bond does not indicate a noticeable change comparing to B3LYP.

Relative sulfonyl radical formation energy has also been computed using B3LYP and M06-2X functionals with the value of  $-10.6$  and  $-37.4$  kJ/mol. Comparing these results

with MP2 shows B3LYP performs poorly by 36.6 kJ/mol error, while M06-2X displays 9.8 kJ/mol error. It is worth mentioning that the applying polar environment ( $\epsilon=78.35$ ) increases calculated relative energy by  $\sim 15$  kJ/mol while increasing the basis set from 6-31G(d) to 6-311G(d) has no effects on the relative calculated energies in both gas and solution phase.

#### 5.3.1.2 C—S(O<sub>2</sub>) Bond Formation

Based on the highest observed spin density on the C adjacent N atom in the initial pyrrolyl radical, the sulfonyl radical obtained from the reaction between S atom in SO<sub>2</sub> and C atom in the pyrrolyl ring has also been studied. In the obtained radical sulfonyl species, the highest spin density using B3LYP and M06-2X functionals is measured for S atom in the -SO<sub>2</sub> group by 0.33 and 0.36 respectively, which agreed with computed value using MP2.

Measured C—S bond length using B3LYP and M06-2X in this species is 1.763 and 1.752 Å respectively, which display an error by 0.025 and 0.036 Å, respectively comparing MP2. Also, the observed changes in  $r(\text{C—S})$  applying the polar environment ( $\epsilon=78.35$ ) is  $< 0.010$  Å. In addition, the larger basis set with triple valence description (6-311G(d)) does not impact the C—S bond length noticeably.

Relative formation energy for this sulfonyl radical molecule has been calculated. Using B3LYP and comparing the obtained result with MP2 represents 25.5 kJ/mol error to  $-50.6$  kJ/mol, while this error decreases to 3.8 kJ/mol using M06-2X to  $-72.3$  kJ/mol. Increasing the measured relative energy resulted in employing the polar environment is  $< 5.4$  kJ/mol. Similarly, no considerable changes were observed with applying the larger basis set of 6-311G(d) in studying the formation of such sulfonyl radical species.

As observed, C—S(O<sub>2</sub>) bond formation by 34.9 (in the gas phase) and 46.7 kJ/mol (in solvation model) is preferred to N—S(O<sub>2</sub>) bond formation using the M06-2X functional.

Since M06-2X shows relatively better performance than B3LYP in calculating the formation energy of C/N—S(O<sub>2</sub>), comparing MP2 functional, this functional is applied as the more accurate and reliable one in calculating the energy in the current study.

### 5.3.1.3 Cross-Linking Reactions

To study the thermodynamically stable products of a radical-based, 3-component reaction involving two equivalents of pyrrolyl and one equivalent of sulfur dioxide, the formation energy of three possible products of diaryl sulfone, sulfonamide, and sulfamide in both gas phase and aqueous solution, using M06-2X functional are outlined in **Table 5.1**.

**Table 5.1.** Calculated formation energy (kJ/mol) of SO<sub>2</sub>-containing compounds using pyrrolyl.

	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;">   <b>A<sub>1</sub></b> </div> <div style="text-align: center;">   <b>B<sub>1</sub></b> </div> <div style="text-align: center;">   <b>C<sub>1</sub></b> </div> </div>
$\Delta E_g$	<div style="display: flex; justify-content: space-around;"> <span>-356.0</span> <span>-341.9</span> <span>-318.2</span> </div>
$\Delta E_{\text{SCRF}}$	<div style="display: flex; justify-content: space-around;"> <span>-350.1</span> <span>-325.2</span> <span>-290.3</span> </div>

As it is clear in **Table 5.1**, formation of diaryl sulfone (**A<sub>1</sub>**), a result of cross-linking of C—S(O<sub>2</sub>)—C is the thermodynamically favored product among sulfonamide (**B<sub>1</sub>**) and sulfamide (**C<sub>1</sub>**) by 14.1 and 37.8 kJ/mol respectively, in the gas phase, which aqueous solution increases the values by 24.9 and 59.8 kJ/mol. Due to the highest observed spin density of the initial pyrrolyl radical on the C atom adjacent N atom and preference of the C—S bond formation to S—N bond, the formation of diaryl sulfone species as the most preferred one among other possible products satisfied our expectation.

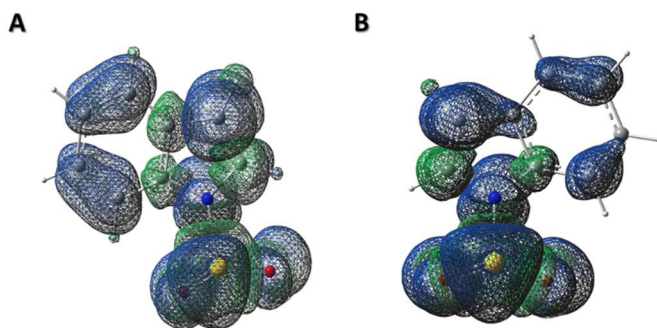
### 5.3.2. Indolyl

The spin density of the indolyl has been calculated using B3LYP functional which represents the highest ones on the C(3) and N of pyrrolyl ring with the value of 0.51 and 0.49, respectively. Conversely, the highest spin density on Ph-ring is related to C(7) with the value of 0.13 (see **Scheme 5.2**).

#### 5.3.2.1 N–S(O<sub>2</sub>) Bond Formation

Employing B3LYP and M06-2X functionals, the formation of the sulfonyl radical species from the reaction of N atom of indolyl and sulfur dioxide has been examined. Calculated spin density using M06-2X indicates that S atom bears the largest spin density of 0.35 which is in a good agreement with MP2 and followed by two O atoms of sulfur dioxide and then N atoms. In addition, comparing the N–S(O<sub>2</sub>) bond lengths using M06-2X and MP2 makes a good compromise in both gas and solvation model with the error < 0.008 Å. It is while the calculated spin contamination in this such open-shell system,  $\langle S^2 \rangle$ , using MP2 functional is 0.05 larger than expectation value (0.75).

Furthermore, calculated spin density using B3LYP functional represents quite equal spin density distribution (~0.21) among S, and N and C(3) atoms of indolyl, not predominantly on the -SO<sub>2</sub> group in the gas phase (**Figure 5.1**).



**Figure 5.1.** Spin density map for indole sulfonyl species using (A) B3LYP and (B) M06-2X functionals, respectively.

Furthermore, applying B3LYP functional, the N—S(O<sub>2</sub>) bond length is elongated by 0.101 Å to 1.815 Å relative to MP2 in the gas phase, while N···S distance in aqueous solution is measured by 2.413 Å with the error of 0.705 Å comparing MP2. Improving the basis set from 6-31G(d) to 6-311G(d) is also caused S···N to lengthen by 0.018 and 0.006 Å in gas and solution, respectively compared to B3LYP/6-31G(d). To consider the influence of long-range interaction on the  $r(\text{S}\cdots\text{N})$  in this species, dispersion interaction effect using B3LYP-D3 employed which shows shortening S···N distance by 0.006 and 0.079 Å to 1.809 and 2.334 Å in gas and aqueous media, respectively.

Due to the observed results related to B3LYP and B3LYP-D3 methods and their inefficiency in describing the most accurate geometry, also the observed spin contamination applying MP2, the relative formation energy for the corresponding sulfonyl radical molecules have been calculated using M06-2X functional, with the value of −35.5 kJ/mol. IEFPCM solvation model increases the energy by 15.6 to −19.9 kJ/mol. The formation energy relative to the corresponding bond formation in our smallest system (pyrrolyl) does not represent a significant change (~2 kJ/mol).

#### 5.3.2.2 C—S(O<sub>2</sub>) Bond Formation

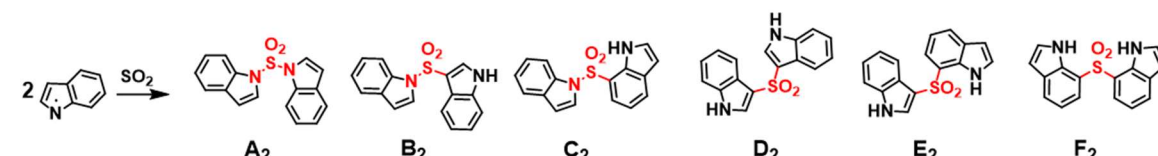
**C(3)—S(O<sub>2</sub>) Formation:** Bond formation between C(3), the atom with the highest spin density in the indolyl, and S atom has examined. Similar to what observed in pyrrolyl ring, SO<sub>2</sub> group shows the highest spin distribution, mostly on the S atom. All B3LYP, M06-2X, and MP2 agree on the obtained results. Using B3LYP and M06-2X functionals,  $r(\text{C—S})$  is measured by 1.764 and 1.749 Å, respectively with the < 0.009 Å error comparing MP2 (despite observed spin contamination). The polar environment also causes  $r(\text{C—S})$  to feel shortening by 0.012 Å using M06-2X and B3LYP/6-31G(d) functionals. In comparison with pyrrolyl system and C—S(O<sub>2</sub>) formation, the C(3)—S(O<sub>2</sub>) formation energy in such model using M06-2X has been raised by 3.5 to −68.8 kJ/mol in gas and decreased by 4.6 kJ/mol to −73.5 kJ/mol in solvation model.

**C(7)—S(O<sub>2</sub>) Formation:** C(7)—S bond formation in Ph-ring has been also investigated (see **Scheme 5.2**). In the obtained sulfonyl species,  $r(\text{C}(7)\text{—S})$  is measured to be longer by approximate 0.028 and 0.036 Å in the gas phase and solvation model, respectively using B3LYP and M06-2X functionals comparing to  $r(\text{C}(3)\text{—S})$ . In addition, C(7)—SO<sub>2</sub> formation energy, using M06-2X has been decreased by 10.5 to −79.3 kJ/mol in the gas phase and increased only by 4.4 to −69.1 kJ/mol in aqueous solution relative to C(3)—SO<sub>2</sub> bond formation.

### 5.3.2.3 Cross-Linking Reactions

Comparing the stability of SO<sub>2</sub>-containing compounds from a 3-components radical reaction of two equivalent of initial indolyl and sulfur dioxide molecules, through the formation energy of the most possible products based on the spin density distribution have been shown in **Table 5.2**.

**Table 5.2.** Calculated formation energy (kJ/mol) of SO<sub>2</sub>-containing compounds using indolyl.

		<b>A<sub>2</sub></b>	<b>B<sub>2</sub></b>	<b>C<sub>2</sub></b>	<b>D<sub>2</sub></b>	<b>E<sub>2</sub></b>	<b>F<sub>2</sub></b>
$\Delta E_g$		-305.5	-337.8	-343.0	-345.6	-359.3	-363.6
$\Delta E_{\text{SCRF}}$		-277.1	-329.0	-320.4	-356.3	-357.6	-348.2

As **Table 5.2** shows three different diaryl sulfones species (compounds **D<sub>2</sub>–F<sub>2</sub>**) are the most thermodynamically stable products than possible diaryl sulfonamides (compounds **B<sub>2</sub>–C<sub>2</sub>**) and sulfamide (**A<sub>2</sub>**), while sulfamide is least favored formations thermodynamically. Interestingly, diaryl sulfone **F<sub>2</sub>**, that is the result of cross-linking of C(7)—S(O<sub>2</sub>)—C(7) is the most preferred one. It is while C(7) on the Ph-ring bears much lower spin density relative to C(3) and N of the indolyl despite the preference of

C(7)—S(O<sub>2</sub>) formation to N—S(O<sub>2</sub>) and C(3)—S(O<sub>2</sub>). It is worth mentioning that diaryl sulfone **D<sub>2</sub>**, the result of cross-linking of C(3)—S(O<sub>2</sub>)—C(3) in which C(3) shows the highest spin density in indolyl species, is 18.0 kJ/mol higher in energy than diaryl sulfone **F<sub>2</sub>**. Obviously, the aqueous solution increases the computed formation energy from 1.7 (in diaryl sulfone **E<sub>2</sub>**) to 28.4 kJ/mol (in diaryl sulfamide **A<sub>2</sub>**) except for diaryl sulfone **D<sub>2</sub>**, which is lowered by 10.7 kJ/mol.

### 5.3.3 Cabazolyl

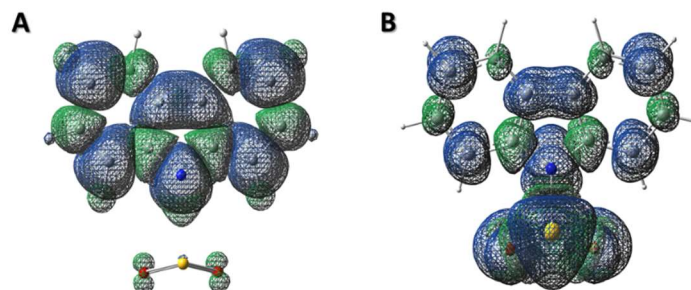
Unlike pyrrolyl and indolyl molecules, in carbazolyl, the most spin density is localized on the N atom with the value of 0.51 calculated which is ~2.5 times more than observed ones on C(1)(3)(6)(8) with the spin density ~0.21 (see **Scheme 5.2**).

#### 5.3.3.1 N—S(O<sub>2</sub>) Bond Formation

N—S(O<sub>2</sub>) bond formation in sulfonyl radical species have been studied using B3LYP, B3LYP-D3, and M06-2X functionals. Results using M06-2X show that S atom in this species has the highest value of spin density (0.34) and the second and third places are related to the two O and N atoms with the value of 0.20 and 0.18, respectively. The  $r(\text{N—S})$  bond in this species is measured 1.723 in the gas phase, which shows an increase of 0.023 and 0.017 Å comparing to the corresponding bond in the pyrrolyl and indolyl species, respectively. The polar environment also increases the  $r(\text{N—S})$  by 0.005 Å.

B3LYP functional has also been used to examine this species. The S···N distance is measured by 2.654 and 2.542 Å in the gas and solution phases, respectively. It is while the observed spin density delocalization value is quite similar to the initial carbazolyl radical with no spin density distribution on SO<sub>2</sub> species (see **Figure 5.2**). Additionally, the observed spin contamination for this species seems to be larger than the expected value; 0.78. Use of Grimme's correction, B3LYP-D3, has shortened the  $r(\text{S}···\text{N})$  by 0.118 and 0.110 Å to 2.536 and 2.432 Å in the gas and solvation phases, respectively.





**Figure 5.2.** Spin density map for carbazole sulfonyl species using (A) B3LYP and (B) M06-2X functionals, respectively.

In spite of the most observed spin density on the N atom of carbazolyl species, using M06-2X functional, the sulfonyl formation energy using has decreased only by 1.8 kJ/mol to  $-39.2$  kJ/mol compared to the corresponding bond formation in pyrrolyl system. Although the polar environment increases this energy by 3.4 kJ/mol comparing pyrrolyl system to  $-18.8$  kJ/mol. It is worth mentioning that the corrected formation energy with BSSE for carbazolyl-(N)···S(O<sub>2</sub>) complex using B3LYP functional is calculated by  $-21.6$  kJ/mol.

### 5.3.3.2 C—S(O<sub>2</sub>) Bond Formation

C—S bond formation between C(1) and C(3), with the highest spin density after N atom in carbazolyl, and S atom of sulfur dioxide have also been examined. In both formed sulfonyl species, -SO<sub>2</sub> group displays the most spin densities with no spin density distribution on the other atoms in carbazole. Using M06-2X and B3LYP functionals, C(1)—S bond length shows an excellent compromise with  $r(\text{C}(7)\text{—S})$  in the indole sulfonyl species. Although C(1)—S(O<sub>2</sub>) bond formation energy has been increased by 7.9 and 6.0 kJ/mol to  $-71.4$  and  $-63.1$  kJ/mol in the gas and aqueous medium, respectively compared to corresponding C(7)—S(O<sub>2</sub>) bond in the indole sulfonyl species.

In addition, examining the C(3)—S(O<sub>2</sub>) bond formation using B3LYP and M06-2X functionals results in the elongation of  $r(\text{C—S})$  by around 0.011 Å relative to the

$r(\text{C}(1)\text{—S})$ . Also, the comparison shows that  $\text{C}(3)\text{—S}$  bond formation energy has been raised by 5.8 to  $-65.6$  kJ/mol relative to  $\text{C}(1)\text{—S}$  bond formation in the gas phase, although the polar environment, decreases it to  $-69.6$  kJ/mol.

Based on the obtained findings in the carbazolyl despite observing the highest spin density distribution on the N atom, the formation of  $\text{N—S}(\text{O}_2)$  formation is not favored over  $\text{C—S}(\text{O}_2)$ .

### 5.3.3.3 Cross-Linking Reactions

Thermodynamic stability comparison among possible SO<sub>2</sub>-containing compounds result from the radical-based 3-component reaction of carbazolyl and sulfur dioxide are listed in **Table 5.3**.

As shown, similar to the cross-linking reaction in indolyl and sulfur dioxide species, three diaryl sulfones species (compound **D<sub>3</sub>–F<sub>3</sub>**) are the thermodynamically preferred products to possible diaryl sulfonamides (compound **B<sub>3</sub>–C<sub>3</sub>**). Not surprisingly among obtained diaryl sulfones, the one which is the result of the cross-linking reaction of  $\text{C}(1)\text{—SO}_2\text{—C}(1)$ , diaryl sulfone **D<sub>3</sub>**, should be considered as the most thermodynamically favored sulfone species by 7.9 and 18.0 kJ/mol relative to diaryl sulfone **E<sub>3</sub>** and **F<sub>3</sub>**, respectively in the gas phase. It should be mentioned that sulfamide (compound **A<sub>3</sub>**), a result of the cross-linking reaction of two N-center radical carbazolyl and SO<sub>2</sub> is the least thermodynamically preferred product, although the N atom in the carbazolyl displays the highest value of spin density distribution.

Employing a polar environment increases the energy of formation from 1.7 (in diaryl sulfone **E<sub>3</sub>**) to 26.6 kJ/mol (in diaryl sulfamide **A<sub>3</sub>**). The only exception occurs for diaryl sulfone **F<sub>3</sub>**, where the SO<sub>2</sub> moiety links two C(3) centers of initial cabazolyl. In this case  $\Delta E_{\text{SCRF}}$  is lower than  $\Delta E_{\text{g}}$  by 11.5 kJ/mol

**Table 5.3.** Calculated formation energy (kJ/mol) of SO<sub>2</sub>-containing compounds applying carbazoyl.

	A <sub>3</sub>	B <sub>3</sub>	C <sub>3</sub>	D <sub>3</sub>	E <sub>3</sub>	F <sub>3</sub>
$\Delta E_g$	-300.1	-330.4	-326.6	-354.9	-347.0	-336.9
$\Delta E_{\text{SCRF}}$	-273.5	-310.4	-317.9	-341.3	-345.3	-348.4

#### 5.3.3.4 Substitution Effects

To study the effects of various substitution on the C—S(O<sub>2</sub>) and N—S(O<sub>2</sub>) bond formation, and cross-linking reactions, H on -C(3) of carbazoyl species is replaced with a number of electrons withdrawing and donating groups (see **Scheme 5.2**).

**3-Methyl Carbazoyl:** As findings show the presence of the -CH<sub>3</sub> group on the -C(3) of carbazole does not indicate the significant changes in the measured N—S(O<sub>2</sub>), C(1)—S(O<sub>2</sub>), C(6)—S(O<sub>2</sub>), and C(8)—S(O<sub>2</sub>) distances relative to the corresponding distances in the unsubstituted carbazoyl species. Although it leads to increasing the formation energy of N—S(O<sub>2</sub>) by 8.1 to -31.1 kJ/mol using M06-2X functional in the gas phase with respect to the analogous bond formation in the carbazoyl species. Additionally, -CH<sub>3</sub> group results in increasing the C—S(O<sub>2</sub>) formation energy by ~3 kJ/mol relative to the corresponding bond formation in the unsubstituted carbazoyl.

**3-Fluoro Carbazoyl:** Similarly, inserting the electron-withdrawing group of -F on the -C(3) of carbazoyl species causes a insignificant changes in the measured N—S(O<sub>2</sub>), C(1)—S(O<sub>2</sub>), C(6)—S(O<sub>2</sub>), and C(8)—S(O<sub>2</sub>) distances relative to analogous bonds in the

unsubstituted carbazolyl. It is while the energy for the formation of N—S(O<sub>2</sub>) has been raised by 11.1 to −28.1 kJ/mol relative to the corresponding bond in the carbazolyl. Also, C—S(O<sub>2</sub>) formation experiences an increase of 6.4–11.6 kJ/mol in comparison with corresponding unsubstituted carbazolyl.

**3-Amino Carbazolyl:** A noticeable influence of -NH<sub>2</sub> substitution firstly is observed in the spin density distribution on the carbazolyl species comparing -CH<sub>3</sub> and -F. The presence such an electron-donating group causes lowering the spin density distribution on the N, C(6), and C(8) to 0.42, 0.11, and 0.13, respectively.

In addition, -NH<sub>2</sub> leads to significant lengthening of N···S(O<sub>2</sub>) distance. Applying the M06-2X functional, the measured  $r(\text{N} \cdots \text{S}(\text{O}_2))$  in the obtained complex result of 3-amino carbazolyl and SO<sub>2</sub> is 2.413 and 2.042 Å in the gas and aqueous environment, respectively. Although these observed distances have been decreased by 0.157 and 0.285 Å relative to measured distances applying B3LYP functional in the gas and solution phase, respectively.

Another considerable impact of -NH<sub>2</sub> presence is observed in C(1)—S(O<sub>2</sub>), C(6)—S(O<sub>2</sub>), and C(8)—S(O<sub>2</sub>) bond formation in which the formation energy has been raised by < ~6 and ~38 kJ/mol in comparison with unsubstituted carbazolyl species in the gas phase and aqueous solution, respectively.

#### 5.3.3.5 Cross-Linking Reactions

Influence of -CH<sub>3</sub> and -F substituents on the radical-based 3-component reaction is also studied (**Table 5.4**). Similar to what resulted of the unsubstituted carbazolyl in the cross-linking reactions, diaryl sulfamide, and sulfones, respectively is considered as the least and most favored products thermodynamically. Among obtained sulfone species (compounds **E4–J4**), products result of the cross-linking of C(1)—SO<sub>2</sub>—C(8) and C(8)—SO<sub>2</sub>—C(8), (diaryl sulfone **F4** and **H4**, respectively) are the preferred thermodynamic ones in which

both C(1) and C(8) have indicated the highest spin density after N atom in the initial substituted carbazole species.

Obviously, despite electron-withdrawing nature of -F relative to the neutral nature of -CH<sub>3</sub>, noticeable changes in the formation energy of SO<sub>2</sub>-containing compounds derived from substituted carbazolyl were not observed (< ~13 kJ/mol).

Expectedly, applying the polar environment increases the formation of the desired compounds. The only exception is observed for diaryl sulfone **J<sub>4</sub>**, a result of C(6)—SO<sub>2</sub>—C(6) cross-linking reaction in which the aqueous medium decreases the energy (see **Table 5.4**).

As earlier discussed, a result of not observing the covalent bond formation between N atom of 3-amino carbazolyl and S atom of sulfur dioxide, the formation of only six possible diaryl sulfone species were examined for the cross-linking 3-component reaction derived from 3-amino carbazolyl species. As **Table 5.4** outlines, expectedly based on the spin density distribution on 3-amino carbazolyl, diaryl sulfones **F<sub>4</sub>** and **H<sub>4</sub>**, obtained from C(1)—SO<sub>2</sub>—C(8) and C(8)—SO<sub>2</sub>—C(8) cross-linking respectively, are regarded as the most thermodynamically stable ones among other products. It should be stated that -NH<sub>2</sub> substitution leads to increasing the formation energy of the related diaryl sulfones *via* cross-linking reactions by < ~39 and < ~45 kJ/mol relative to -F and -CH<sub>3</sub> substitutions in the gas phase. Considerably, the polar environment raises the observed difference though by less than ~68 kJ/mol.

#### 5.3.4 Radical Xiamycin

Spin density distribution on the xiamycin radical species as a bulky alkyl-substituted carbazolyl ring has also been examined. Results show that inserting such a bulky alkyl substituent in carbazolyl species does not affect the calculated spin density noticeably comparing to the initial unsubstituted carbazolyl.

**Table 5.4.** Calculated formation energy (kJ/mol) of SO<sub>2</sub>-containing compounds from substituted carbazoyl.

$2 \text{ (substituted carbazoyl)} \xrightarrow{\text{SO}_2}$

$\text{X} = \text{CH}_3, \text{F}, \text{NH}_2$

		A <sub>4</sub>	B <sub>4</sub>	C <sub>4</sub>	D <sub>4</sub>	E <sub>4</sub>	F <sub>4</sub>	G <sub>4</sub>	H <sub>4</sub>	I <sub>4</sub>	J <sub>4</sub>
$\Delta E_g$	X=CH <sub>3</sub>	-294.8	-324.6	-323.4	-320.2	-343.2	-347.4	-340.0	-343.4	-338.2	-329.9
	X=F	-292.8	-315.1	-319.1	-316.2	-329.8	-338.9	-331.4	-338.5	-334.0	-324.6
	X=NH <sub>2</sub>					-298.5	-302.7	-296.6	-300.2	-294.8	-287.0
$\Delta E_{\text{SCRF}}$	X=CH <sub>3</sub>	-266.8	-302.6	-302.6	-308.9	-327.1	-330.1	-334.6	-327.2	-335.5	-338.2
	X=F	-266.7	-296.1	-300.8	-309.8	-316.0	-325.6	-331.2	-328.4	-337.2	-340.0
	X=NH <sub>2</sub>					-259.3	-263.0	-268.1	-261.1	-269.7	-272.2

5.3.4.1 N–S(O<sub>2</sub>) Bond Formation

Applying three different functionals including B3LYP, B3LYP-D3, and M06-2X, the formation of N–S(O<sub>2</sub>) was examined. Comparing the results with the corresponding species in the unsubstituted carbazoyl ring does not display considerable changes in the measured  $r(\text{N}–\text{S}(\text{O}_2))$  using M06-2X;  $< 0.006 \text{ \AA}$  in the gas and solution phase. Similarly,

it should be noted both B3LYP and B3LYP-D3 functionals lengthen the  $r(\text{N}\cdots\text{S}(\text{O}_2))$  distance to 2.643 and 2.523 Å, respectively in the gas phase. Although the polar environment decreases the distances by 0.127 and 0.117 Å to 2.516 and 2.406 Å, respectively.

One of the effects of the presence of this bulky substituent is observed in calculating N—S(O<sub>2</sub>) formation energy. It shows an increase of 20.2 kJ/mol relative to the analogous bond formation in unsubstituted carbazoyl system to −19.0 kJ/mol in the gas phase. Although employing the polar environment decreases the observed difference to 3.8 kJ/mol. It's worth mentioning that the corrected formation energy with BSSE for (N)⋯S(O<sub>2</sub>) complex in this species using B3LYP functional is calculated by −12.6 kJ/mol, an increase of 9.0 kJ/mol relative to carbazoyl-N⋯S(O<sub>2</sub>) complexes without any substituent.

#### 5.3.4.2 C—S(O<sub>2</sub>) Bond Formation

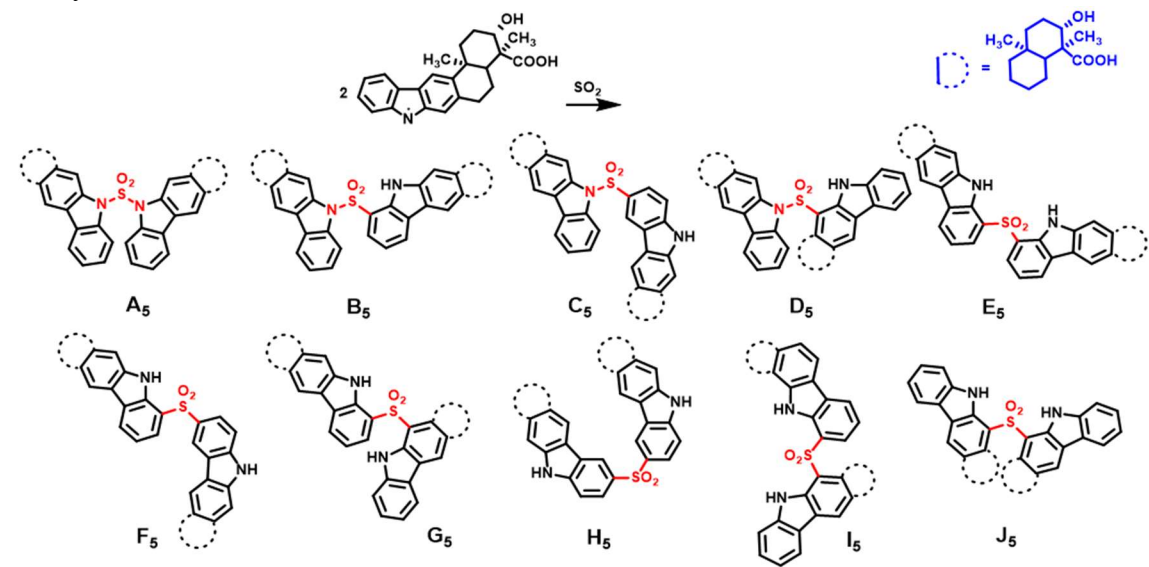
Additionally, influence such bulky substitution on the C—S(O<sub>2</sub>) bond formation applying two B3LYP and M06-2X functionals were studied. Comparison C(1)—S(O<sub>2</sub>) and C(5)—S(O<sub>2</sub>) with corresponding ones in the unsubstituted carbazoyl system shows insignificant changes in the measured bond lengths. The only exception is related to the  $r(\text{C}(12)\text{—S}(\text{O}_2))$  which indicates a 0.017 and 0.021 Å bond lengthening using B3LYP and M06-2X, respectively comparing the unsubstituted carbazoyl system. Result of the observed bond lengthening, the related bond formation energy increases by 8.8 and 10.9 kJ/mol to −62.6 and −52.2 kJ/mol in the gas and polar environment, respectively.

#### 5.3.4.3 Cross-Linking Reactions

Plausible products resulted from the cross-linking reaction of radical xiamicin and SO<sub>2</sub> are listed in **Table 5.5**. Expectedly, based on the results from previous smaller systems, diaryl sulfamide **A5** and sulfone **F5**, a result of the cross-linking of C(1)—SO<sub>2</sub>—C(5), are regarded

as the least and most thermodynamically stable products, respectively. This finding is in good agreement with spin density distribution on the less substituted Ph-ring of radical xiamycin.

**Table 5.5.** Calculated formation energy (kJ/mol) of SO<sub>2</sub>-containing compounds from xiamycin.



	A <sub>5</sub>	B <sub>5</sub>	C <sub>5</sub>	D <sub>5</sub>	E <sub>5</sub>	F <sub>5</sub>	G <sub>5</sub>	H <sub>5</sub>	I <sub>5</sub>	J <sub>5</sub>
$\Delta E_g$	-297.9	-310.1	-326.6	-325.5	-329.7	-374.5	-344.6	-317.3	-341.1	-315.7
$\Delta E_{\text{SCRF}}$	-280.6	-298.8	-321.7	-308.8	-321.6	-387.9	-327.1	-355.8	-359.4	-310.7

As expected, applying the polar environment increases the calculated formation energy by ~5-17 kJ/mol. But some exceptions are observed in the cross-linking reaction of C(1)—S(O<sub>2</sub>)—C(5) (sulfone **F<sub>5</sub>**), C(5)—SO<sub>2</sub>—C(5) (sulfone **H<sub>5</sub>**), and C(12)—S(O<sub>2</sub>)—C(5) (sulfone **I<sub>5</sub>**) in which the polar environment decreases the formation of energy by ~13-39 kJ/mol.

Interestingly, diaryl sulfone **H<sub>5</sub>** and **I<sub>5</sub>** are the same ones which were identified as the bioproducts of a bacterial synthesis of some sulfa compounds antibiotics by flavoenzyme *XiaH* and sulfur dioxide capturing<sup>35</sup> and the other as mentioned is the most thermodynamically stable one in the gas phase (compound **F<sub>5</sub>**). Also, among three possible



diaryl sulfonamides species (compound **B5**, **C5**, and **D5**), diaryl sulfonamide **C5** is the most thermodynamically feasible one in both gas and polar environment. Surprisingly, this compound is the other observed bioproduct of flavoenzyme catalyzed bacterial synthesis.

## 5.4 Conclusions

SO<sub>2</sub>-containing compounds, including sulfones (R<sub>3</sub>C-SO<sub>2</sub>-CR<sub>3</sub>), sulfonamides (R<sub>3</sub>C-SO<sub>2</sub>-NR<sub>2</sub>), and sulfamides (R<sub>2</sub>N-SO<sub>2</sub>-NR<sub>2</sub>) are known as highly efficient versatile species with a wide spectrum of chemical and biological activities and properties. Some DFT and conventional *ab initio* methods have been applied to study the formation of such diaryl SO<sub>2</sub>-containing moieties through a radical-based three-component reaction involving two equivalents of a number of small to bulky N-containing heterocyclic compounds and one equivalent of sulfur dioxide. Our findings show that C—S(O<sub>2</sub>) bond formation is preferred to N—S(O<sub>2</sub>) bond formation which necessarily does not agree with the observed spin density distribution in the initial radical N-containing heterocyclic species.

Preference of C—S(O<sub>2</sub>) to N—S(O<sub>2</sub>) results in the formation of diaryl sulfones (R<sub>3</sub>C-SO<sub>2</sub>-CR<sub>3</sub>) and diaryl sulfamides (R<sub>2</sub>N-SO<sub>2</sub>-NR<sub>2</sub>) as the most and least thermodynamically favored possible SO<sub>2</sub>-containing compounds in the current study, respectively.

In addition, the influence of the different substitutions on the formation of the desired SO<sub>2</sub>-containing compounds has been investigated. The results indicate that presence of -CH<sub>3</sub> and -F substitutions as the neutral and electron-withdrawing groups, respectively on one of our medium-sized, N-containing heterocyclic radical (carbazolyl) do not considerably impact the obtained results compared with the unsubstituted species. Although the observed changes with inserting electron-donating group such as -NH<sub>2</sub> are considerable both structurally and energetically.

More significantly, the obtained results related to the initial radical xiamycin and studying the most thermodynamically preferred SO<sub>2</sub>-containing compounds are in a good agreement with experimentally observed diaryl sulfonamide and sulfones antibiotics which

were identified as the bioproducts through a bacterial synthesis of flavoenzyme (*XiaH*) and sulfur dioxide capturing.

As well, in the current study the ability of some DFT methods and conventional *ab initio* in the formation of SO<sub>2</sub>-containing compounds of interest were assessed. Based on our finding, the M06-2X functional as the *meta*-GGA functional, among other applied DFT methods, B3LYP and B3LYP-D3, performs well in predicting geometrical parameters and energies, while B3LYP and B3LYP-D3 tend to overestimate the C—S(O<sub>2</sub>) and N—S(O<sub>2</sub>) bond lengths and significantly affect the related structures and formation energies.

On the other hand, observing the larger spin contamination than expected value using *ab initio* MP2 by increasing the size of the models of interest causes bonds lengthening and changes in the energy compared with other functionals.

### 5.5 References

1. Stipanuk, M. H., Sulfur Amino Acid Metabolism: Pathways for Production and Removal of Homocysteine and Cysteine. *Annu. Rev. Nutr.* **2004**, *24*, 539-577.
2. Du, S. X.; Jin, H. F.; Bu, D. F.; Zhao, X.; Geng, B.; Tang, C. S.; Du, J. B., Endogenously Generated Sulfur Dioxide and its Vasorelaxant Effect in Rats. *Acta Pharm. Sinic.* **2008**, *29*, 923-930.
3. Wang, X. B.; Jin, H. F.; Tang, C. S.; Du, J. B., Significance of Endogenous Sulphur-Containing Gases in the Cardiovascular System. *Clin. Exp. Pharmacol. P.* **2010**, *37*, 745-752.
4. Hong, T., Advances in the Study on Endogenous Sulfur Dioxide in the Cardiovascular System. *Chin. Med. J.* **2014**, *127*, 3803-3807.
5. Mathew, N. D.; Schlipalius, D. I.; Ebert, P. R., Sulfurous Gases as Biological Messengers and Toxins: Comparative Genetics of Their Metabolism in Model Organisms. *J. Toxicol.* **2011**, *2011*, 1-15.

6. Li, R. J.; Meng, Z. Q.; Xie, J. F., Effects of sulfur Dioxide Derivatives on Four Asthma-Related Gene Expressions in Human Bronchial Epithelial Cells. *Toxicol. Lett.* **2007**, *175*, 71-81.
7. Min, J. Y.; Min, K. B.; Cho, S. I.; Paek, D. Y., Combined Effect of Cigarette Smoking and Sulfur Dioxide on Heart Rate Variability. *Int. J. Cardiol.* **2009**, *133*, 119-121.
8. Wang, W. Y.; Ji, X. Y.; Du, Z. M.; Wang, B. H., Sulfur Dioxide Prodrugs: Triggered Release of SO<sub>2</sub> via a Click Reaction. *Chem. Commun.* **2017**, *53*, 1370-1373.
9. Wang, X. B.; Du, J. B.; Cui, H., Sulfur dioxide, A Double-Faced Molecule in Mammals. *Life Sci.* **2014**, *98*, 63-67.
10. Emmett, E. J.; Willis, M. C., The Development and Application of Sulfur Dioxide Surrogates in Synthetic Organic Chemistry. *Asian J. Org. Chem.* **2015**, *4*, 602-611.
11. M- J. El-Hibri; Weinberg, S. A., Polysulfones. In *Encyclopedia of Polymer Science and Technology*, Wiley: New York, 2002.
12. Bissereet, P.; Blanchard, N., Taming Sulfur Dioxide: A Breakthrough for its Wide Utilization in Chemistry and Biology. *Org. Biomol Chem.* **2013**, *11*, 5393-5398.
13. Goadsby, P. J.; Ferrari, M. D.; Stovner, L. J.; Senard, J. M.; Jackson, N. C.; Poole, P. H.; Stat, C., Eletriptan in Acute Migraine: A Double-Blind, Placebo-Controlled Comparison to Sumatriptan - Reply. *Neurology* **2000**, *55*, 736-736.
14. Le, Y.; Ji, H.; Chen, J. F.; Shen, Z. G.; Yun, J.; Pu, M., Nanosized Bicalutamide and its Molecular Structure in Solvents. *Int. J. Pharm.* **2009**, *370*, 175-180.
15. Firke, S. D.; Bari, S. B., Synthesis, Biological Evaluation and Docking Study of Maleimide Derivatives Bearing Benzenesulfonamide as Selective COX-2 Inhibitors and Anti-Inflammatory Agents. *Biorg. Med. Chem.* **2015**, *23*, 5273-5281.
16. Ilardi, E. A.; Vitaku, E.; Njardarson, J. T., Data-Mining for Sulfur and Fluorine: An Evaluation of Pharmaceuticals to Reveal Opportunities for Drug Design and Discovery. *J. Med. Chem.* **2014**, *57*, 2832-2842.

17. Nissinen, L.; Ojala, M.; Langen, B.; Dost, R.; Pihlavisto, M.; Kapyla, J.; Marjamäki, A.; Heino, J., Sulfonamide Inhibitors of  $\alpha 2\beta 1$  Integrin Reveal the Essential Role of Collagen Receptors in in vivo Models of Inflammation. *Pharmacol. Res. Perspect.* **2015**, *3*, e00146.
18. Gangapuram, M.; Mazzio, E.; Eyunni, S.; Soliman, K. F. A.; Redda, K. K., Synthesis and Biological Evaluation of Substituted N-[3-(1H-Pyrrol-1-yl)methyl]-1,2,5,6-tetrahydropyridin-1-yl]benzamide/benzene Sulfonamides as Anti-Inflammatory Agents. *Arch. Pharm.* **2014**, *347*, 360-369.
19. Zhao, C.; Rakesh, K. P.; Ravidar, L.; Fang, W. Y.; Qin, H. L., Pharmaceutical and Medicinal Significance Of Sulfur (S<sup>VI</sup>)-Containing Motifs for Drug Discovery: A Critical Review. *Eur. J. Med. Chem.* **2019**, *162*, 679-734.
20. Liu, N. W.; Liang, S.; Manolikakes, G., Recent Advances in the Synthesis of Sulfones. *Synthesis-Stuttgart* **2016**, *48*, 1939-1973.
21. Trost, B. M.; Ghadiri, M. R., Sulfones As Chemical Chameleons - Cyclization Via 1,1-Dipole Synthons. *J. Am. Chem. Soc.* **1984**, *106*, 7260-7261.
22. Abrunhosa, I.; Gulea, M.; Masson, S., Efficient New Protocol to Synthesize Aromatic and Heteroaromatic Dithioesters. *Synthesis-Stuttgart* **2004**, 928-934.
23. Gui, J. H.; Zhou, Q. H.; Pan, C. M.; Yabe, Y.; Burns, A. C.; Collins, M. R.; Ornelas, M. A.; Ishihara, Y.; Baran, P. S., C-H Methylation of Heteroarenes Inspired by Radical SAM Methyl Transferase. *J. Am. Chem. Soc.* **2014**, *136*, 4853-4856.
24. Greig, I. R.; Coste, E.; Ralston, S. H.; van 't Hof, R. J., Development of Triarylsulfonamides as Novel Anti-Inflammatory Agents. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 816-820.
25. Capasso, C.; Supuran, C. T., Sulfa and Trimethoprim-Like Drugs - Antimetabolites Acting as Carbonic Anhydrase, Dihydropteroate Synthase and Dihydrofolate Reductase Inhibitors. *J. Enzyme Inhib. Med. Chem.* **2014**, *29*, 379-387.

26. Supuran, C. T., Carbonic Anhydrase Inhibition and the Management of Hypoxic Tumors. *Metabolites* **2017**, *7*, 48.
27. Carta, F.; Mannelli, L. D.; Pinard, M.; Ghelardini, C.; Scozzafava, A.; McKenna, R.; Supuran, C. T., A Class of Sulfonamide Carbonic Anhydrase Inhibitors with Neuropathic Pain Modulating Effects. *Biorg. Med. Chem.* **2015**, *23*, 1828-1840.
28. Scozzafava, A.; Supuran, C. T.; Carta, F., Antiobesity Carbonic Anhydrase Inhibitors: A Literature and Patent Review. *Expert Opin. Ther. Pat.* **2013**, *23*, 725-735.
29. Anusha, S.; Sinha, A.; Rajeev, C. P. B.; Chu, T. T. T.; Mathai, J.; Huang, X. M.; Fuchs, J. E.; Shivananju, N.; Bender, A.; Preiser, P. R.; Rangappa, K. S.; Basappa; Chandramohanadas, R., Synthesis, Characterization and in Vitro Evaluation of Novel Enantiomerically-Pure Sulphonamide Antimalarials. *Org. Biomol Chem.* **2015**, *13*, 10681-10690.
30. Mujumdar, P.; Teruya, K.; Tonissen, K. F.; Vullo, D.; Supuran, C. T.; Peat, T. S.; Poulsen, S.-A., An Unusual Natural Product Primary Sulfonamide: Synthesis, Carbonic Anhydrase Inhibition, and Protein X-ray Structures of Psammaphin C. *J. Med. Chem.* **2016**, *59*, 5462-5470.
31. Liu, G.; Fan, C. B.; Wu, J., Fixation of Sulfur Dioxide into Small Molecules. *Org. Biomol Chem.* **2015**, *13*, 1592-1599.
32. Zheng, D. Q.; An, Y. Y.; Li, Z. H.; Wu, J., Metal-Free Aminosulfonylation of Aryldiazonium Tetrafluoroborates with DABCO Center.(SO<sub>2</sub>)<sub>2</sub> and Hydrazines. *Angew. Chem. Int. Ed.* **2014**, *53*, 2451-2454.
33. Li, Y. W.; Mao, R. Y.; Wu, J., N-Radical Initiated Aminosulfonylation of Unactivated C(sp<sup>3</sup>)-H Bond through Insertion of Sulfur Dioxide. *Org. Lett.* **2017**, *19*, 4472-4475.
34. Li, W. F.; Beller, M.; Wu, X. F., Catalytic Conversion of Aryl Triazenes into Aryl Sulfonamides Using Sulfur Dioxide as the Sulfonyl Source. *Chem. Commun.* **2014**, *50*, 9513-9516.

35. Baunach, M.; Ding, L.; Willing, K.; Hertweck, C., Bacterial Synthesis of Unusual Sulfonamide and Sulfone Antibiotics by Flavoenzyme-Mediated Sulfur Dioxide Capture. *Angew. Chem. Int. Ed.* **2015**, *54*, 13279-13283.
36. de Lima Procópio, R. E.; da Silva, I. R.; Martins, M. K.; de Azevedo, J. L.; de Araújo, J. M., Antibiotics Produced by Streptomyces. *Braz. J. Infect. Dis.* **2012**, *16*, 466-471.
37. Teufel, R.; Agarwal, V.; Moore, B. S., Unusual Flavoenzyme Catalysis in Marine Bacteria. *Curr. Opin. Chem. Biol.* **2016**, *31*, 31-39.
38. Young, N. A., Main Group Coordination Chemistry at Low Temperatures: A Review of Matrix Isolated Group 12 to Group 18 Complexes. *Coord. Chem. Rev.* **2013**, *257*, 956-1010.
39. Cui, G. K.; Zhang, F. T.; Zhou, X. Y.; Huang, Y. J.; Xuan, X. P.; Wang, J. J., Acylamido-Based Anion-Functionalized Ionic Liquids for Efficient SO<sub>2</sub> Capture through Multiple-Site Interactions. *ACS Sustainable Chem. Eng.* **2015**, *3*, 2264-2270.
40. Kurten, T.; Lane, J. R.; Jorgensen, S.; Kjaergaard, H. G., A Computational Study of the Oxidation of SO<sub>2</sub> to SO<sub>3</sub> by Gas-Phase Organic Oxidants. *J. Phys. Chem. A* **2011**, *115*, 8669-8681.
41. Steudel, R.; Steudel, Y., Charge-Transfer Complexes between the Sulfur Molecules SO<sub>2</sub>, S<sub>2</sub>O, S<sub>3</sub>, SONH, and SOCl<sub>2</sub> and the Amine Donors NH<sub>3</sub> and NMe<sub>3</sub> - A Theoretical Study. *Eur. J. Inorg. Chem.* **2007**, 4385-4392.
42. Keller, J. W., Sulfur Dioxide-Pyridine Dimer. FTIR and Theoretical Evidence for a Low-Symmetry Structure. *J. Phys. Chem. A* **2015**, *119*, 10390-10398.
43. Markovic, D.; Volla, C. M. R.; Vogel, P.; Varela-Alvarez, A.; Sordo, J. A., BCl<sub>3</sub>-Mediated Ene Reaction of Sulfur Dioxide and Unfunctionalized Alkenes. *Chem. Eur. J.* **2010**, *16*, 5969-5975.

44. Krupa, J.; Wierzejewska, M., Structural and Spectroscopic Properties of Complexes Formed between HNCS and SO<sub>2</sub> in Low Temperature Matrices. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2017**, *183*, 144-149.
45. Sinha, S.; Raj, A.; Al Shoaibi, A. S.; Chung, S. H., Reaction Mechanism for m-Xylene Oxidation in the Claus Process by Sulfur Dioxide. *J. Phys. Chem. A* **2015**, *119*, 9889-9900.
46. Sarojini, K.; Krishnan, H.; Kanakam, C. C.; Muthu, S., Synthesis, structural, spectroscopic studies, NBO analysis, NLO and HOMO-LUMO of 4-methyl-N-(3-nitrophenyl)benzene Sulfonamide with Experimental and Theoretical Approaches. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2013**, *108*, 159-170.
47. Chandran, A.; Mary, Y. S.; Varghese, H. T.; Panicker, C. Y.; Pazdera, P.; Rajendran, G., FT-IR, FT-Raman Spectroscopy And Computational Study of (E)-4-((anthracen-9-ylmethylene)amino)-N-carbamimidoylbenzene Sulfonamide. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2011**, *79*, 1584-1592.
48. Murthy, P. K.; Suneetha, V.; Armakovic, S.; Armakovic, S. J.; Suchetan, P. A.; Giri, L.; Rao, R. S., Synthesis, Characterization and computational Study of the Newly Synthetized Sulfonamide Molecule. *J. Mol. Struct.* **2018**, *1153*, 212-229.
49. Govindarasu, K.; Kavitha, E.; Sundaraganesan, N., Synthesis, Structural, Spectral (FTIR, FT-Raman, UV, NMR), NBO and First Order Hyperpolarizability Analysis of N-Phenylbenzenesulfonamide by Density Functional Theory. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2014**, *133*, 417-431.
50. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.;

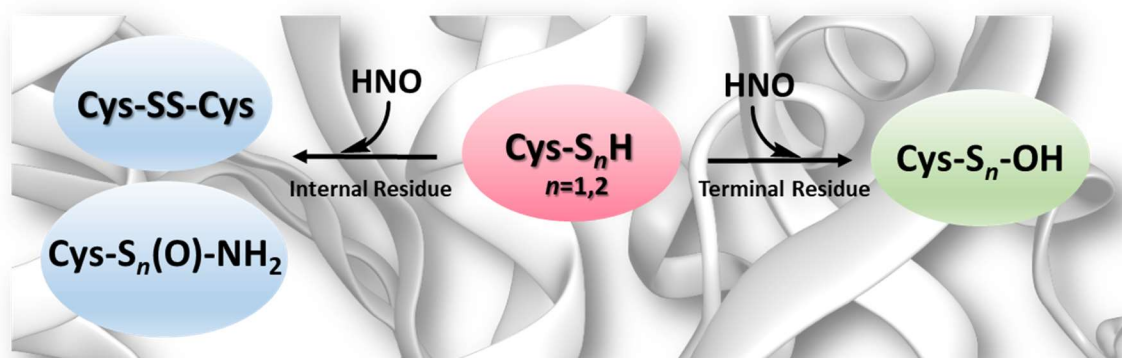
Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; N. Rega; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; J. V. Ortiz; Cioslowski, J.; Fox, D. J. *Gaussian 09*, E.01; Gaussian, Inc.: Wallingford CT, 2015.

51. Becke, A. D., A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372-1377.
52. Lee, C. T.; Yang, W. T.; Parr, R. G., Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. *Phys Rev B* **1988**, *37*, 785-789.
53. Zhao, Y.; Truhlar, D. G., The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215-241.
54. Grimme, S., Density Functional Theory with London Dispersion Corrections. *WIREs. Comput. Mol. Sci.* **2011**, *1*, 211-228.
55. Boys, S. F.; Bernardi, F., The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. *Mol. Phys.* **2002**, *100*, 65-73.



## CHAPTER 6

# Computational Study on HNO-Derived Post-Translational Modifications of Cysteinyll and Cysteinyll Persulfide



## 6.1 Introduction

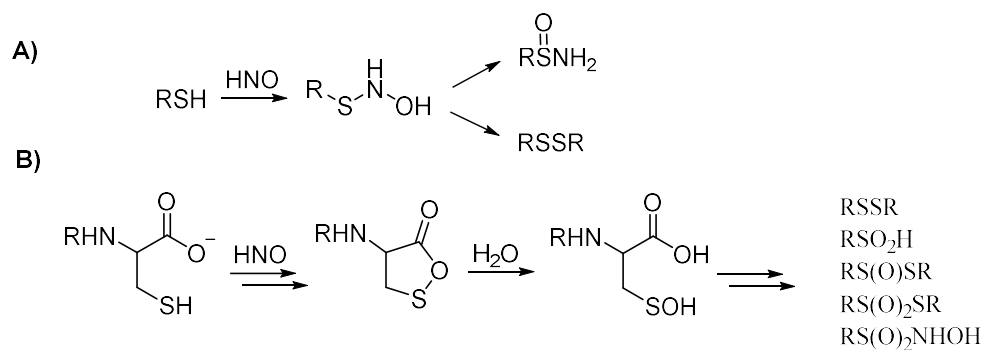
Post-translational modification (PTM) of proteins (*e.g.*, glycosylation, ubiquitination, sumoylation) is important for a variety of physiological processes including, for example, signaling and degradation, as well as expanding the chemical functionality of proteins and enzymes.<sup>1-2</sup> Indeed, it has been estimated that up to two thirds of all proteins expressed in eukaryotes are at least temporarily phosphorylated.<sup>3-5</sup> As a result, understanding the formation of such modifications, their effects, and roles has long been of considerable interest.<sup>6-7</sup> Within peptides, however, a number of residues such as seriny, threony, and tyrosyl are known to more commonly be the target of PTMs, *e.g.*, phosphorylation. However, cysteiny is also known to be able to undergo a range of post-translational modifications under oxidative conditions. For example, they can form disulfide bonds, the most common covalent crosslink found in proteins, persulfides, sulfonamides and oxidized derivatives such as sulfenic acid.<sup>8-10</sup>

Recently, it has been shown that cysteine can be modified by reaction with nitroxyl (Azanone: HNO). The latter is the protonated one-electron reduced form of nitric oxide and is now known to be a potent species in biological systems exhibiting many physiological effects including enzyme inhibition and regulation.<sup>11-13</sup> In addition, it has also shown potential pharmacological significance as a therapeutic agent for a variety of disease including heart failure, cancer, alcoholism, and vascular relaxation.<sup>13-20</sup> Notably, the biological activity of HNO has been postulated to result in large part from its ability to modify thiol groups, in particular, that of cysteine.<sup>21-23</sup>

It was commonly held that depending on the thiol concentration, the thiols reaction with HNO leads to formation of N-hydroxysulfenamide (RSNHOH) which then reacts further to produce disulfide (RSSR) and/or sulfinamide (RS(O)NH<sub>2</sub>) (**Scheme 6.1A**).<sup>24-29</sup> Indeed, such induced disulfide or sulfinamide modifications have been observed in several proteins including aldehyde dehydrogenase (ALDH), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), papain, and yeast transcription factor ACE1.<sup>17, 30-35</sup> However, it is now known

that the reaction of HNO with peptide thiols is influenced by other factors including the environment's hydrophobicity/polarity, the thiol's  $pK_a$ , and nearby acid/base functional groups. For example, under standard conditions HNO reacts with thiols with a rate of  $k = 3 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ .<sup>36</sup> However, with the active site thiol GAPDH it reacts at an even higher rate of  $>1 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$  due to a nearby histidyl acting as a base to facilitate their reaction.<sup>31, 37</sup> Recently, Keceli *et al.* experimentally observed that reaction of HNO with C-terminus cysteiny's resulted in the formation of sulfohydroxamic acid  $\text{RS(O}_2\text{)NHOH}$ , thiosulfonate ( $\text{RS(O}_2\text{)SR}$ ) and thiosulfinate ( $\text{RS(O)SR}$ ) as well as disulfides and sulfenamide.<sup>38</sup> They proposed that, as shown in **Scheme 6.1B**, the availability of the adjacent carboxylate led *via* a cyclic intermediate to sulfenic acid, which then reacts further to give the broader range of products.

**Scheme 6.1.** Schematic illustration of the reaction and resulting products of HNO with a: (A) thiol, and (B) C-terminus cysteiny thiol.



There have been several computational investigations on the reactions of HNO with S-containing compounds.<sup>36, 39-42</sup> For instance, Sherman *et al.* used B3LYP, MP2 and CBS-QB3 methods, in combination with a polarized continuum (IEFPCM) solvation model to examine mechanisms by which HNO may react with a variety of free thiol derivatives.<sup>41</sup> In particular, they concluded that in hydrophobic environments (*i.e.*, low polarity) the kinetic production of disulfides was preferred while thiols with electron-withdrawing substituents, and in the presence of suitable bases, preferentially and irreversibly formed

sulfenamides. However, the reactions examined did not consider the impact of a cysteinyI's position in a peptide, or the energetic bio-feasibility of those pathways. More recently, Ivanova *et al.* used several DFT methods to examine the related process of HNO formation *via* S-thiolation involving thiols and S-nitrosothiols (RSNO).<sup>42</sup> They concluded that such reactions proceed *via* a zwitterionic intermediate,  $\text{RSS}^+(\text{R})\text{N}(\text{H})\text{O}^-$  followed by S—N bond dissociation to give HNO and the RSSR disulfide. They also noted that the reaction was facilitated in an aqueous environment *via* water-assisted proton shuttling.

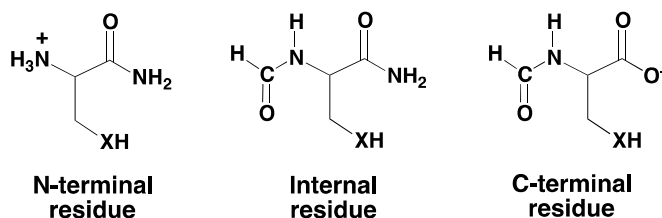
Recently, S-sulfhydration of cysteine residues (*i.e.*, formation of the persulfide derivative  $\text{RSS}^-/\text{H}$ ) has been recognized as important to a variety of physiological processes from redox signaling to enzyme regulation and thus consequently is of increasing interest.<sup>43-45</sup> Due to the inherent instability and high reactivity of such persulfide species, their chemical biology is not well-known. However, their lower  $\text{p}K_a$  and greater nucleophilicity compared to thiols, has led to the proposal that they may more readily react with HNO within biological environments. Unfortunately, to date, there have been few or no experimental or computational studies on the reaction of HNO with cysteinyI persulfides.

In this present study, we have used DFT-based methods in combination with a systematic series of chemical models to examine the reaction of HNO with the cysteinyI and cysteinyI persulfide residues and the effect thereon of their position and/or environment in a peptide or protein. More specifically, the reactions of HNO with C- and N-terminus, as well as peptide-internal cysteinyI residues, and cysteinyI persulfides have been examined. The effect and role of the presence of carboxylic acid (C-terminus) and amine (N-terminus) functional groups, and the polarity of the environment have also been considered.

## 6.2 Computational Methods

All calculations were performed using the Gaussian 09 suite of programs.<sup>46</sup> Optimized geometries were obtained out using the density functional theory (DFT) *meta*-GGA functional M06-2X, in conjunction with the 6-311G(d,p) basis set. This level of theory has previously been shown to provide reliable optimized geometries of sulfur-containing molecules.<sup>47-48</sup> In addition, the polarity of the surrounding environment was included by use of the IEFPCM solvation method. Dielectric constants ( $\epsilon$ ) of 4, 10, and 78.35 were used to model a non- or polar protein environment, and an aqueous solution environment respectively, and as has been commonly used in computational enzymology studies.<sup>49</sup> That is, optimized geometries were obtained at the IEFPCM( $\epsilon = Y$ )-M06-2X/6-311G(d,p) level of theory. Harmonic vibrational frequencies were calculated at the same level of theory for all optimized structures to ensure they were minima or mechanism-relevant first-order transition structures. Relative energies were obtained by performing single point calculations at the IEFPCM( $\epsilon = Y$ )-M06-2X/6-311+G(2df,p) level of theory on the above optimized structures; IEFPCM( $\epsilon = Y$ )-M06-2X/6-311+G(2df,p)// IEFPCM( $\epsilon = Y$ )-M06-2X/6-311G(d,p). Topological analyses of the densities were performed using the AIMALL program.<sup>50</sup> To model cysteiny and cysteiny persulfide residues in the different possible positions in which they may occur within a peptide, without biases that may arise from use of a specific X-ray crystal peptide's structure, the structures of the residue shown in **Scheme 6.2** were used.

**Scheme 6.2.** Schematic illustration of the chemical models used for the cysteiny (X = S) and cysteiny persulfide (X = S-S) residues within (internal) and in the C- and N-terminus positions, of a peptide.



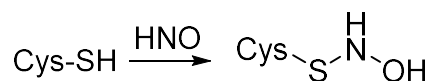
## 6.3 Results and Discussion

### 6.3.1 Reaction of HNO and CysteinyI Residue

#### 6.3.1.1 N-hydroxysulfenamide Formation

The expected product of the reaction of HNO and the cysteinyI residue is N-hydroxysulfenamide intermediate. As shown in **Scheme 6.3**, the formation of S—N bond and transfer of the S—H proton of cysteine to the O atom of HNO are the key processes in this reaction.

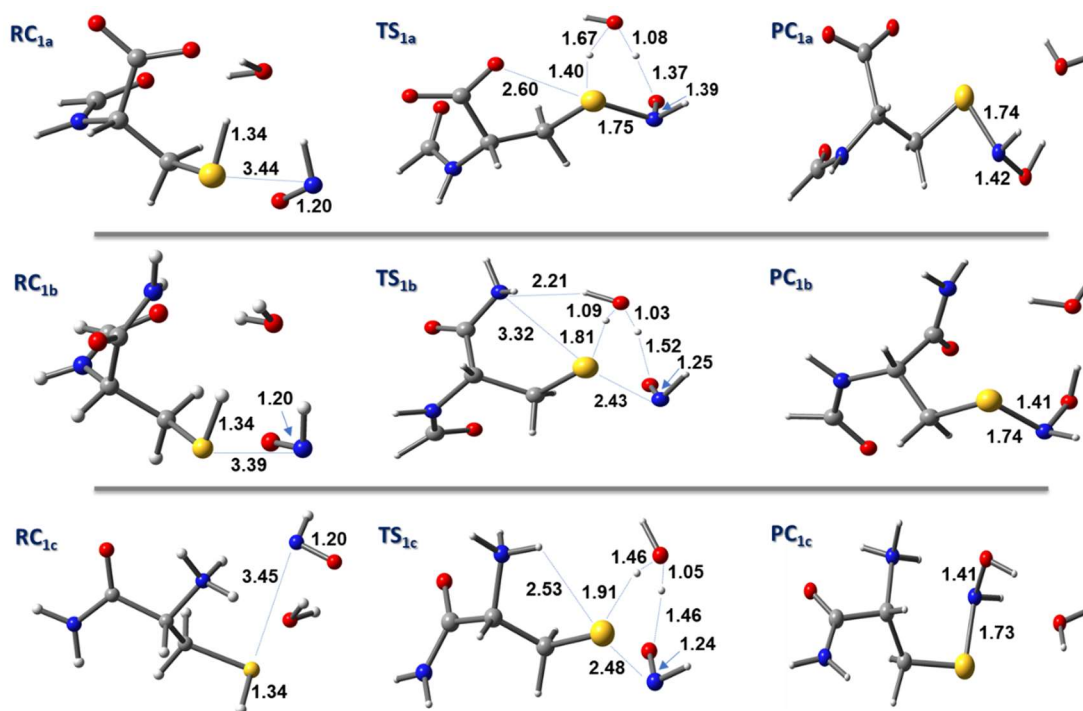
**Scheme 6.3.** Schematic representation of N-hydroxysulfenamide formation.



Considering the cysteine residue position and its plausible effects on the reactivity of HNO toward S atom of cysteine, this reaction has been studied using three different models of cysteine including C-terminus, internal, and N-terminus residues.

In the aqueous media, explicit water molecules potentially assist proton transferring from S to O atom through a 6-membered cyclic transition structure. **Figure 6.1** represents the obtained optimized reactant complexes, transition structures, and product complexes of models of our interest. Obviously, comparison of **RC<sub>1</sub>** and **PC<sub>1</sub>** applying different models of cysteine residue does not indicate the considerable changes in the measured  $r(\text{S—H})$  and  $r(\text{CysS} \cdots \text{N}_{\text{HNO}})$  in the reactant complexes,  $r(\text{S—N})$  and  $r(\text{N—O})$  in the product complexes, respectively. However, the most significant changes are observed in the obtained transition structures.

Comparing the transition structures shows  $r(\text{S—H})$  of cysteine residue (from 1.34 Å) is significantly lengthened to 1.40 Å in C-terminus, 1.81 Å in internal and 1.91 Å in N-terminus cysteine residues. The  $\text{S} \cdots \text{N}$  bond distance also has been measured by 1.75 Å in C-terminus, 2.43 and 2.48 Å in internal and N-terminus cysteine residues, respectively.

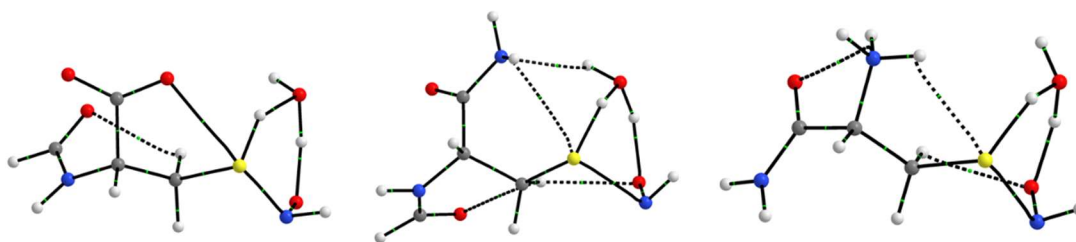


**Figure 6.1.** Optimized reactant complexes ( $RC_1$ ), transition structures ( $TS_1$ ), and product complexes ( $PC_1$ ) with selected bond distances in ångströms (Å) of the water-catalyzed reaction of HNO with C-terminus, internal, and N-terminus cysteine residues from up to down, respectively.

Additionally, the molecular graphs for the studied transition structures using QTAIM are shown in **Figure 6.2**. The QTAIM calculated values of electron density ( $\rho$ ) and Laplacian of the electron density ( $\nabla^2\rho$ ) at the bond critical point (BCP) for the  $S\cdots O_{COO^-}$  in the obtained transition structure using C-terminus model of cysteine residue are 0.027 and 0.087 a.u., respectively which indicate the closed-shell interaction type between S and O atom of the carboxylate group. Interestingly, these values lie within the ranges for  $\rho$  (0.002-0.035 a.u.) and  $\nabla^2\rho$  (0.024-0.139 a.u.) previously suggested as a typical H-bonding.<sup>51-52</sup> It should be mentioned such this closed-shell interaction between  $S\cdots O_{COO^-}$  is replaced with a weaker closed-shelled interaction of  $S\cdots N_{NH_2}$  and  $S\cdots H_{NH_3^+}$  in the internal and N-terminus cysteines, respectively (see **Figure 6.2**).

The calculated parameters of QTAIM analysis related to C-terminus cysteine residue reveal that  $S \cdots N_{HNO}$  interaction in the obtained transition structure with  $\rho$  and  $\nabla^2\rho$  of 0.200 and  $-0.354$  a.u., respectively, is characterized as the covalent bond which is consistent with the measured distance of  $1.75 \text{ \AA}$ . It is while the QTAIM calculated values of  $\rho$  and  $\nabla^2\rho$  at the BCPs of  $S \cdots N_{HNO}$  interaction in the transition structures modeled with internal and N-terminus cysteine residue are 0.051 and 0.086, and 0.045 and 0.085 a.u. respectively.

It is worth stating that a closed-shell interaction type between  $HNOO \cdots H_{CH_2}$  with  $\rho$  and  $\nabla^2\rho$  value of 0.014 and 0.048, and 0.010 and 0.039 a.u. respectively, at the BCPs is observed for internal and N-terminus models of cysteine residue in their obtained transition structures in the reaction with HNO. This type of interaction is not characterized in the transition structure obtained for the reaction of HNO with C-terminus cysteiny.



**Figure 6.2.** QTAIM molecular graphs of the obtained transition structures of the water-catalyzed reaction of HNO with C-terminus, internal, and N-terminus cysteine residues from left to right, respectively.

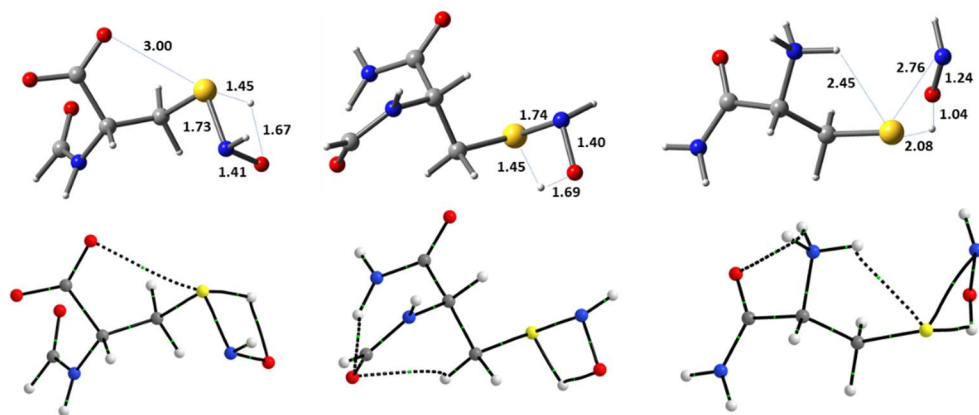
In addition to studying the water-assisted proton transferring mechanism, the unassisted proton transferring process from S to  $O_{HNO}$  also has been examined. The obtained 4-membered cyclic transition structures and the related QTAIM molecular graphs are shown in **Figure 6.3**. The calculated values of  $\rho$  and  $\nabla^2\rho$  at the BCP for S and O atom of carboxylate in the model of C-terminus cysteine indicates a weaker closed-shell interaction with the  $\rho$  and  $\nabla^2\rho$  of 0.014 and 0.051 a.u. respectively, comparing the corresponding one in the presence of an explicit water molecule. As **Figure 6.3** displays the closed-shelled interaction of  $S \cdots N_{NH_2}$  has not been characterized in the obtained transition structure using



the internal model of cysteine residue comparison with corresponding water-catalyzed reaction.

Also the values of  $\rho$  and  $\nabla^2\rho$  at the BCP for S and N<sub>HNO</sub> atom interaction are measured by 0.208 and  $-0.390$  a.u. and, 0.178 and  $-0.417$  a.u. for C-terminus and internal cysteine residues, respectively which with geometrical analysis indicate the stronger S—N interaction relative to analogous interaction in the presence of an explicit solvent molecule. It is while for the reaction of the N-terminus model of cysteine residue with HNO, the calculated values of  $\rho$  and  $\nabla^2\rho$  at the BCP for S $\cdots$ N<sub>HNO</sub> interaction are 0.027 and 0.078 a.u. respectively, which implies a closed-shell interaction type in the obtained transition state relative to the covalent interaction of S $\cdots$ N<sub>HNO</sub> in the C-terminus and internal cysteine residues.

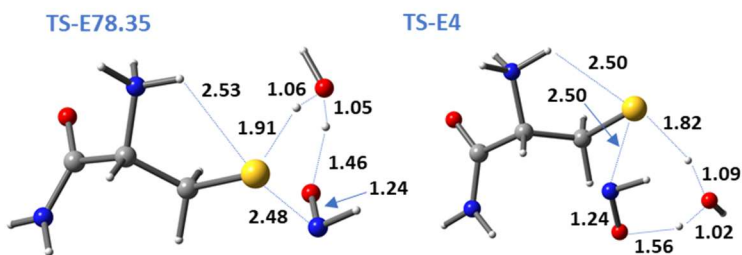
These findings point towards the preference of S—N bond formation to S—H bond breaking in the reaction of C-terminus and internal cysteinyl with HNO in the absence of explicit solvent molecules.



**Figure 6.3.** Optimized transition structures with selected bond distances in ångstroms (Å) and QTAIM molecular graphs of the reaction of HNO with C-terminus, internal, and N-terminus cysteine residues from left to right, respectively.

Considering the influence of the polarity of the environment on the reactivity of HNO toward selected models of cysteiny, the formation of N-hydroxysulfenamide was examined in a different environment with a dielectric constant of 78.35, 10, and 4.

Geometrical parameters of reactants, transitions structures, and product complexes of the reaction between HNO and C-terminus and internal cysteine residue in the presence of an explicit solvent molecule, are not impacted remarkably in environments with different polarities. However, it has been observed that decreasing the polarity from  $\epsilon=78.35$  to  $\epsilon=10$  has also insignificant effects on the geometrical parameters of obtained complexes of water-catalyzed reaction between N-terminus cysteine residue and HNO, applying the polarity of 4 causes a conformational change. As shown in **Figure 6.4**, this observed rotation keeps transferring proton away from positively charged  $\text{NH}_3$  center and assists S—H proton transferring to O atom of HNO.

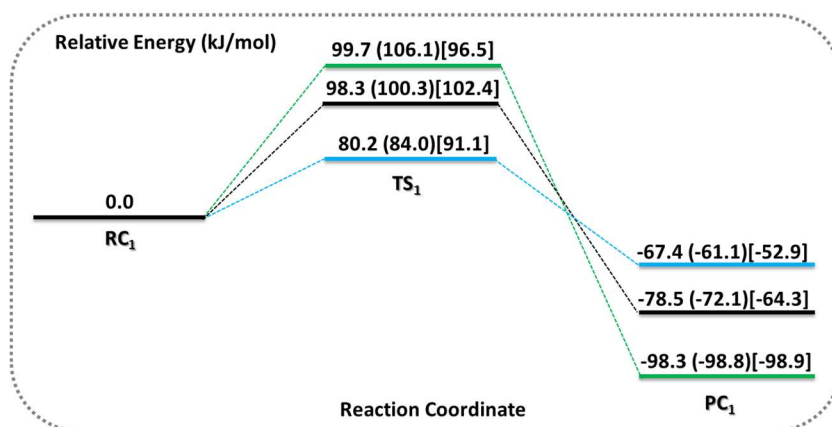


**Figure 6.4.** Optimized 6-membered cyclic transition structures in  $\epsilon=78.35$  and  $\epsilon=4$  with selected bond distances shown in ångströms (Å) for the water-catalyzed reaction of HNO with N-terminus cysteine residue.

The obtained potential energy surface of N-hydroxysulfenamide formation is displayed in **Figure 6.5**. As it is obvious, the calculated activation barriers of 99.7, 98.3, and 80.2 kJ/mol are implied to the reaction of N-terminus, internal, and C-terminus cysteine residues respectively, in the polar environment with a dielectric constant of 78.35. Comparison with the activation barrier of the reaction of HNO and methanethiol (73.5 kJ/mol), the calculated activation barriers indicate an increase of by more than  $\sim 7$  kJ/mol. It is worth mentioning

that inclusion of an explicit water and its contribution in proton transferring process from S to O<sub>HNO</sub> atom results in lowering the activation energy by 77.5, 34.2, and 40.5 kJ/mol in the transition structures modeled with N-terminus, internal, and C-terminus cysteine residues respectively, that states the significant function of the solvent molecule to proceed with reaction.

Furthermore, decreasing the polarity of the solvent from a dielectric constant of 78.35 to 4 leads to raising the activation barrier by 4.1 and 10.9 kJ/mol using internal and C-terminus models of cysteine residues, respectively. Also, the reaction energy increases by ~14 kJ/mol using these two models, concomitant changing the polarity from 78.35 to 4.



**Figure 6.5.** Obtained potential energy surface for the formation of N-hydroxysulfenamide. The green, black and blue colored surface is for the reactivity of HNO toward N-terminus, internal, and C-terminus cysteine residues, respectively. Values in parentheses and brackets are calculated in  $\epsilon=10$  and 4, respectively.

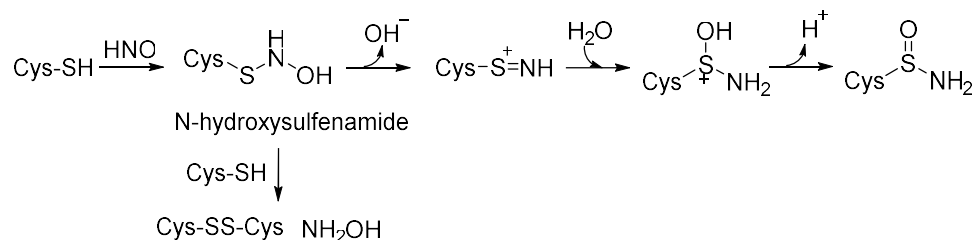
It should be noted that the activation barrier for the reaction of HNO with N-terminus cysteine residue is increased by 6.4 from 99.7 to 106.1 kJ/mol by changing a dielectric constant from 78.35 to 10. It is while applying a dielectric constant of 4 lowers the activation barrier to 96.5 kJ/mol (see **Figure 6.4** for the conformational change). Nevertheless, the reaction energy is not affected significantly by changing the polarity of the environment in this model of cysteine.

Based on the previous computational studies, the  $pK_a$  value of protonated nitrogen atom in N-hydroxysulfenamide, ranges from 11.7 to 15.3.<sup>41</sup> On the other hand, the low  $pK_a$  of protonated O atom of N-hydroxysulfenamide and its decomposition to water and  $RS^+=NH$  species, strongly suggests that the N-protonated form of N-hydroxysulfenamide to be the expected moiety in the solution and physiological conditions.

### 6.3.1.2 Reactions Resulted of Protonated N-Hydroxysulfenamide Derived from Internal Cysteiny and HNO

As stated previously, N-hydroxysulfenamide is assumed to yield the sulfinamide *via* a series of rearrangements and the intermediacy of alkyliminosulfonium species ( $R-S^+=NH$ ) or disulfide moieties (**Scheme 6.4**). In both pathways, the N-protonated form of N-hydroxysulfenamide ( $Cys-SNH_2^+OH$ ) will likely be considered as the intermediate due to its high  $pK_a$  and acting  $HO^-$  and  $NHOH^-$  as the poor leaving groups to form  $Cys-S=NH^+$  or  $Cys-S^+$  species. It should be mentioned that O-protonated form of N-hydroxysulfenamide ( $Cys-SNHOH_2^+$ ) is not found as a stable species and directly decomposes to  $Cys-S^+=NH$  and water molecule, which is in agreement with the low quasi- $pK_a$  of  $RSNHOH_2^+$  species with constrained N—O bond reported by McCulla and coworkers.<sup>41</sup> All these findings specify N atom as the protonation state of N-hydroxysulfenamide in the solution.

**Scheme 6.4.** Schematic representation of reaction of N-hydroxysulfenamide to form sulfinamide and disulfide.

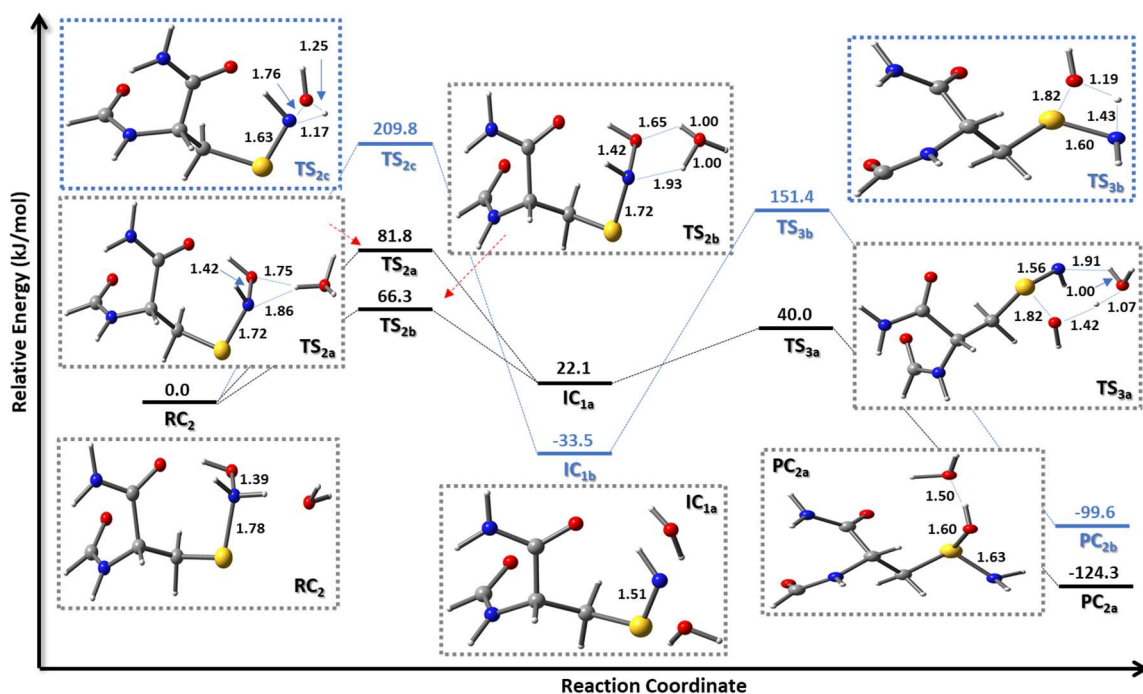


**Sulfonamide.H<sup>+</sup> (Cys-S<sup>+</sup>(OH)-NH<sub>2</sub>) Formation:** The 1,2-H shift from the N-protonated N-hydroxysulfenamide to its O atom leads to the formation of Cys-S<sup>+</sup>=NH species. Acting as proton-transfer catalyst, an explicit water molecule of solvent assists mentioned proton shift either by H-bonding to the transferring proton and dragging the N—H proton to the OH group through a 3-membered cyclic transition structure (**TS<sub>2a</sub>**) at a cost of 81.8 kJ/mol or by simultaneous accepting and donating proton through a 5-membered cyclic transition structure (**TS<sub>2b</sub>**) at a cost of 66.3 kJ/mol relative to the **RC<sub>2</sub>** (see **Figure 6.6**). A considerable geometrical change of the 1,2-H shift is shortening the S—N bond length from 1.78 Å in the **RC<sub>2</sub>** to 1.72 and 1.51 Å in the **TS<sub>2a,2b</sub>**, and Cys-S<sup>+</sup>=NH species (**IC<sub>1</sub>**), respectively. Meanwhile, *r*(N—O) lengthens by 0.03 to 1.42 Å in the **TS<sub>2a,2b</sub>**. NBO charge analysis of S and N atoms of Cys-S<sup>+</sup>=NH has also been calculated by 1.20 and −0.72, respectively which is attributed to more distribution of electron on the N atom in the Cys-S<sup>+</sup>=NH species relative to reactant complex and obtained transition structures. Both obtained transition structures lead to the formation of Cys-S<sup>+</sup>=NH (**IC<sub>1a</sub>**) that lies 22.1 kJ/mol higher in energy than **RC<sub>2</sub>**.

Unassisted 1,2-H shift, in the absence of any explicit solvent molecules, also results in the desired intermediate through a 3-membered cyclic transition structure (**TS<sub>2c</sub>**) with the activation barrier of 209.8 kJ/mol. The noticeable geometrical changes compared to the water-assisted 1,2-H shift are related to the *r*(S—N) which decreased by 0.15 to 1.63 Å and *r*(N—O) that increased significantly by 0.37 from 1.39 to 1.76 Å in the **TS<sub>2c</sub>**.

Subsequently, the positively charged S atom in the obtained Cys-S<sup>+</sup>=NH species is likely to undergo a nucleophilic attack by water molecule to yield O-protonated form of sulfonamide (see **Figure 6.6**). This reaction proceeds *via* a 4-membered cyclic transition structure (**TS<sub>3b</sub>**), with the simultaneous S—O bond formation and proton transferring to N atom which costs 151.4 kJ/mol with respect to the corresponding reactant complex. Although the contribution of explicit water and involving in proton transferring process, through a 6-membered cyclic transition structure (**TS<sub>3a</sub>**) lowers the activation barrier to

40.0 kJ/mol. Obtained  $\text{TS}_{3a}$  results in O-protonated form of sulfinamide ( $\text{PC}_{2a}$ ), that lies 124.3 kJ/mol lower in energy than  $\text{RC}_2$ . The  $r(\text{S—N})$  bond length experiences bond increasing from 1.51 in  $\text{Cys-S}^+=\text{NH}$  to 1.60 and 1.56 Å in the 4-membered and 6-membered cyclic transition structures, respectively and then to 1.63 Å in the product complex ( $\text{PC}_{2a}$ ). As explained, in the rearrangement of N-protonated N-hydroxysulfenamide, the formation of  $\text{Cys-S}^+=\text{NH}$  should be regarded as the rate-determining step.



**Figure 6.6.** Potential energy surface with optimized structures and selected bond distances in ångströms (Å) for the formation of sulfinamide.H<sup>+</sup>. The black and blue colored surfaces are for the water-catalyzed and uncatalyzed reaction.

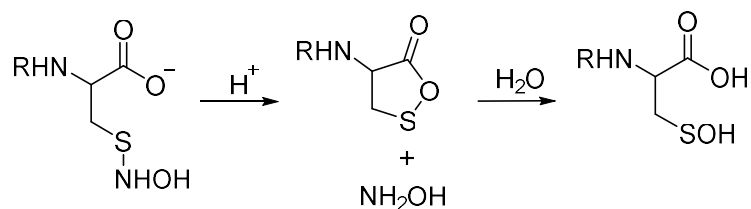
**Disulfide Formation:** Given the high  $pK_a$  of N-protonated N-hydroxysulfenamide, and implausibility of  $\text{NHOH}^-$  reacting as the leaving group for the  $\text{S}_{\text{N}}2$  substitution mechanism, reaction of the neutral cysteine residue ( $\text{Cys-SH}$ ) toward  $\text{Cys-SNH}_2^+\text{OH}$  for the formation of  $\text{Cys-S—S}^+(\text{H})\text{-Cys}$  is considered. The calculated values of  $\Delta E^\ddagger$  and  $\Delta E$  are 225.6 and

42.6 kJ/mol, respectively. Although based on the low  $pK_a$  of  $\text{CH}_3\text{-S-S-CH}_3$  reported by Sherman *et al.*, in the solution<sup>41</sup>, the deprotonation of  $\text{R-SS}^+(\text{H})\text{-R}$  is likely anticipated to form disulfide irreversibly. In this case, the formation of deprotonated  $\text{Cys-S-S-Cys}$  through a 4-membered ring transition structure with simultaneous  $\text{S-S}$  formation and  $\text{H}$  transfer from  $\text{S}$  to  $\text{N}$  atom by the activation barrier of 241.8 kJ/mol forms the neutral disulfide which lies 70.0 kJ/mol lower in energy with respect to the reactant complex. On the other side, in the solution, the activation barrier for  $\text{S}_\text{N}2$  reaction between a neutral nucleophile and positively charged electrophile would be expected to be higher than the gas phase. As a result, for the formation of disulfide, the deprotonated cysteine residue ( $\text{Cys-S}^-$ ) and  $\text{Cys-SNH}_2^+\text{OH}$  are considered as the potential reactants for the substitution reaction to form deprotonated  $\text{Cys-S-S-Cys}$  by  $\Delta E^\ddagger = 122.0$  and  $\Delta E = -132.2$  kJ/mol with the respect to the  $\text{Cys-S}^-$  and  $\text{Cys-SNH}_2^+\text{OH}$  complex.

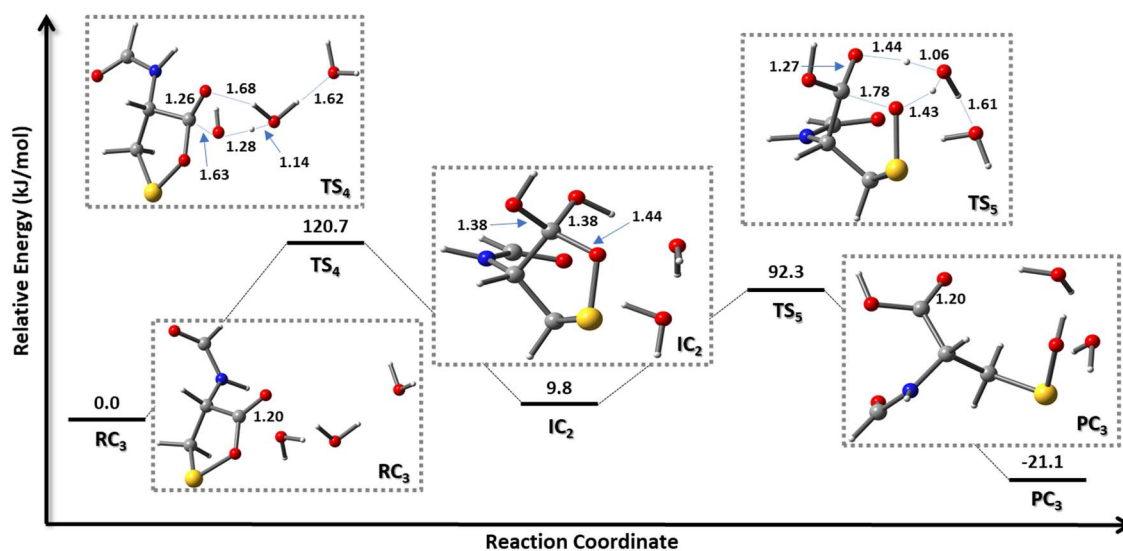
### 6.3.1.3 Reactions Resulted of Protonated N-Hydroxysulfenamide Derived from C-terminus Cysteiny and HNO

Our computational studies have shown that the presence of C-terminus cysteine residue impacts the formation of sulfinamide or disulfide moieties. Interestingly, using C-terminus cysteine residue, no optimized geometries of the N-protonated form of N-hydroxysulfenamide could be located. Instead, a nucleophilic attack to the backside of  $\text{S}$  atom by  $\text{O}$  atom of carboxylate causes the formation of a cyclic structure, a derivative of 1,2-oxathiolane, and hydroxylamine by the  $\Delta E = -1155.6$  kJ/mol (see **Scheme 6.5**). This cyclic species then can be hydrolyzed to form cysteine sulfenic acid.

**Scheme 6.5.** Schematic illustration of cysteine sulfenic acid formation from N-hydroxysulfenamide.



**Figure 6.7** represents the mechanism for the hydrolysis of the obtained cyclic intermediate to cysteine sulfenic acid. As shown, a 2-step water-catalyzed hydrolysis reaction initiates through the nucleophilic attack of hydroxyl water to the electrophilic C of the ester and protonation of the ester carbonyl to form a tetrahedral intermediate (**IC<sub>2</sub>**) with the activation barrier of 120.7 kJ/mol (**TS<sub>4</sub>**). The obtained intermediate undergoes a further reaction to reveal the cysteine sulfenic acid (**PC<sub>3</sub>**) that lies 21.1 kJ/mol lower in energy with respect to the **RC<sub>3</sub>**.



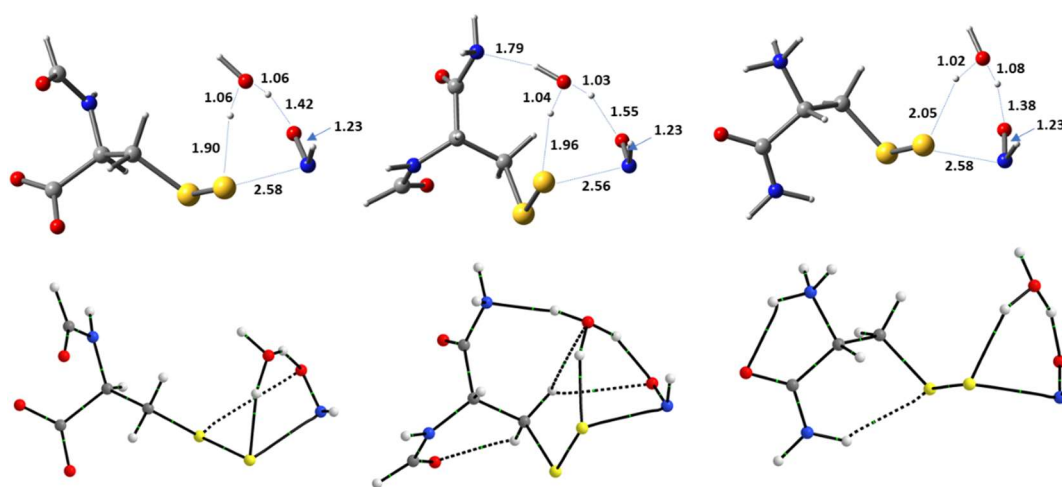
**Figure 6.7.** Potential energy surface of the water-catalyzed sulfenic acid formation with optimized structures and selected bond lengths in ångströms (Å).



### 6.3.2 Reaction of HNO and Cysteiny Persulfide

#### 6.3.2.1 Cys-SS-NHOH Formation

Also, the reactivity of HNO toward the highly nucleophilic terminal sulfur atom in cysteiny persulfide (Cys-SSH) rather than cysteiny has been investigated in this study. Obtained water-assisted 6-membered cyclic transition structures of three different models of cysteiny persulfide including C-terminus, internal and N-terminus cysteiny persulfide with their corresponding QTAIM molecular graphs are displayed in **Figure 6.8**.



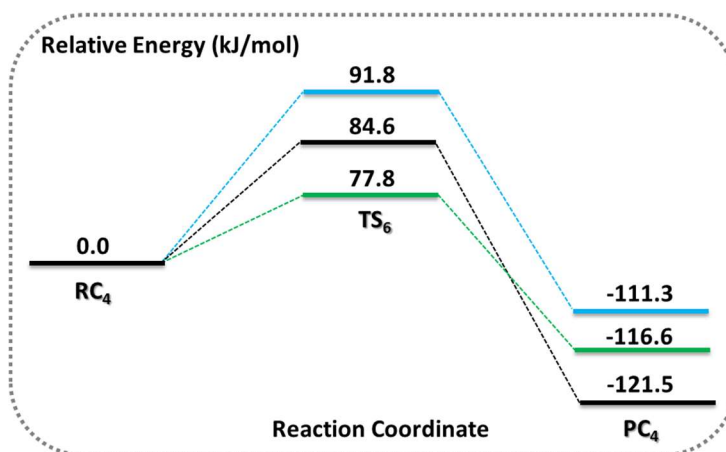
**Figure 6.8.** Optimized transition structures with selected bond distances in ångströms (Å) and QTAIM molecular graphs of the water-catalyzed reaction of HNO with C-terminus, internal, and N-terminus cysteine persulfide residues.

Obviously, no significant changes in  $r(\text{S}\cdots\text{H})$  and  $r(\text{S}\cdots\text{N}_{\text{HNO}})$  in the obtained transition structures have been observed for the solvent-catalyzed formation of Cys-SS-NHOH using various models of cysteine persulfide. In addition, the calculated QTAIM values of  $\rho$  and  $\nabla^2\rho$  at the BCPs related to  $\text{S}\cdots\text{N}_{\text{HNO}}$  in these models imply the closed-shell interaction type between sulfur and nitrogen atoms which should be considered as weaker ones relative to the corresponding  $\text{S}\cdots\text{N}_{\text{HNO}}$  in cysteines models.

Also, **Figure 6.9** represents the potential energy surface resulting from the reaction of various cysteiny persulfide models and HNO. As displayed, higher nucleophilicity of

terminal sulfur of cysteinyl persulfide leads to lowering the activation barrier of Cys-SS-NH-OH formation in the internal and N-terminus cysteinyl persulfide by 13.7 and 21.9 kJ/mol, respectively with the respect to the corresponding Cys-S-NHOH formation in  $\epsilon=78.35$  (compare **Figure 6.9** and **6.5**), despite shorter  $r(\text{S}\cdots\text{N}_{\text{HNO}})$  and stronger closed-shell interaction between S and  $\text{N}_{\text{HNO}}$  atoms in the obtained transition structures for cysteinyl species.

The only exception is observed for the reaction of HNO and C-terminus cysteine persulfide residue, in which the activation barrier has been raised by 11.6 to 91.8 kJ/mol relative to the analogous cysteine residue. This increase in energy might be attributed to the observed contribution of  $\text{COO}^-$  function, especially a closed-shell interaction of  $\text{S}\cdots\text{O}_{\text{COO}^-}$  in the C-terminus cysteinyl to stabilize the transition structure which is not characterized in the transition structure resulted from the reaction of C-terminus cysteinyl persulfide and HNO (compare **Figure 6.2** and **6.8**). Except for the mentioned observation, formation of Cys-SS-NH-OH both kinetically and thermodynamically is preferred to corresponding Cys-S-NHOH species.

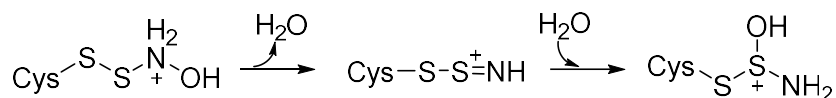


**Figure 6.9.** Obtained potential energy surface for the formation of Cys-SS-NH-OH. The green, black and blue colored surface is for the reactivity of HNO toward N-terminus, internal, and C-terminus cysteinyl persulfide, respectively.

### 6.3.2.2 Reactions Resulted of Cys-SS-NH<sub>2</sub><sup>+</sup>OH Derived from Internal Cysteiny Persulfide and HNO

**Cys-SS<sup>+</sup>(OH)-NH<sub>2</sub> Formation:** Similarly, the formation of Cys-SS<sup>+</sup>(OH)-NH<sub>2</sub>, followed by N-protonation of Cys-SS-NHOH has been investigated (**Scheme 6.6**).

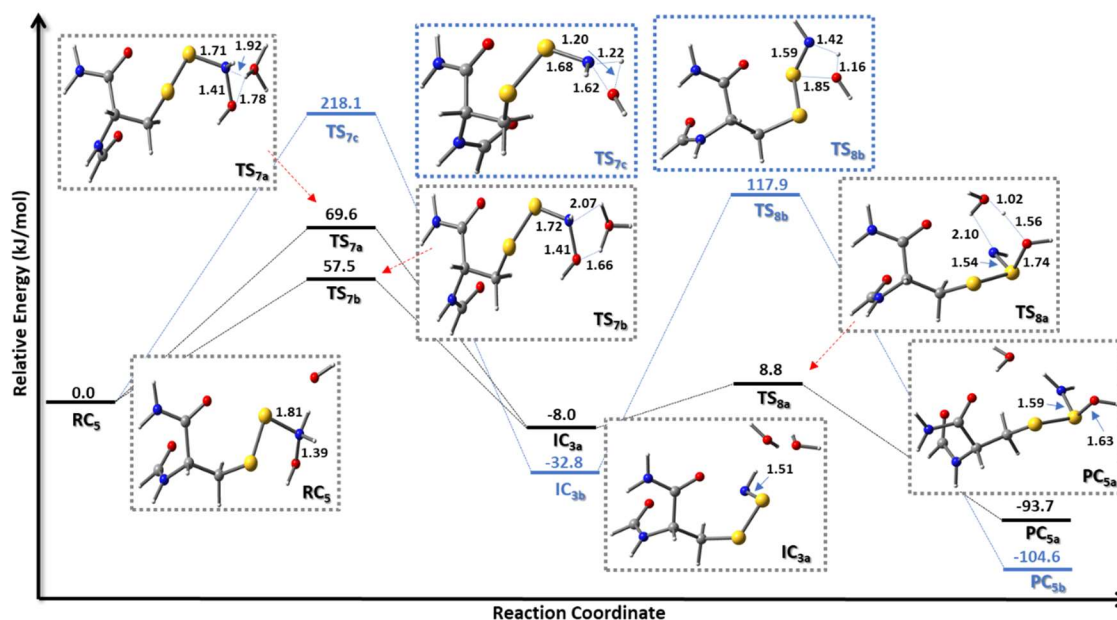
**Scheme 6.6.** Schematic representation of Cys-SS<sup>+</sup>(OH)-NH<sub>2</sub> formation from Cys-SS-NH<sub>2</sub><sup>+</sup>OH.



As **Figure 6.10** indicates a 2-step reaction mechanism of Cys-SS<sup>+</sup>(OH)-NH<sub>2</sub> formation initiates with a 1,2-H shift from N to O atom of Cys-SS-NH<sub>2</sub><sup>+</sup>OH to form Cys-SS<sup>+</sup>=NH species as the intermediate. The solvent-catalyzed reaction proceeds *via* either a 3-membered cyclic transition structure (**TS<sub>7a</sub>**) in which O atom of H<sub>2</sub>O forms H-bonding with transferring proton or a 5-membered cyclic transition structure (**TS<sub>7b</sub>**) with the activation barrier of 69.6 and 57.5 kJ/mol, respectively. Both transition structures lead to Cys-SS<sup>+</sup>=NH formation (**IC<sub>3a</sub>**) that lies 8.0 kJ/mol lower in energy relative to the reactant complex (**RC<sub>5</sub>**). In comparison with Cys-S<sup>+</sup>=NH species, the formation of Cys-SS<sup>+</sup>=NH is considered as the more thermodynamically stable intermediate by 30.1 kJ/mol. Shortening the *r*(S—N) from 1.81 Å in the reactant complex to 1.72 and 1.51 Å in the obtained transition structures and Cys-SS<sup>+</sup>=NH respectively, is the significant geometrical change which is consistent with *r*(N—O) lengthening.

Water-catalyzed nucleophilic attack of a water molecule to the positively charged S atom of the obtained intermediate results in formation of Cys-SS<sup>+</sup>(OH)-NH<sub>2</sub> (**PC<sub>5a</sub>**) with the activation barrier of 8.8 kJ/mol (**TS<sub>8a</sub>**), which shows a decrease of 32.2 kJ/mol relative to the corresponding reaction of Cys-S<sup>+</sup>(OH)-NH<sub>2</sub> formation (compare **Figure 6.6** and **6.10**). As observed, despite the kinetically favored formation of Cys-SS<sup>+</sup>(OH)-NH<sub>2</sub> over Cys-S<sup>+</sup>(OH)-NH<sub>2</sub> species, it is not regarded as the thermodynamic production comparing

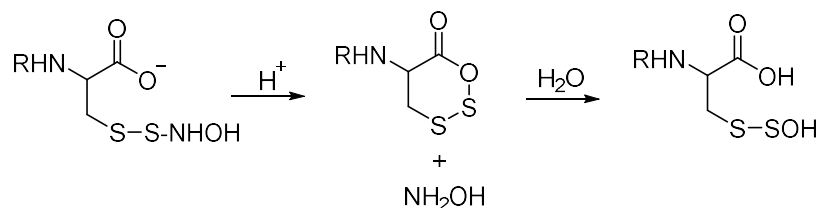
Cys-S<sup>+</sup>(OH)-NH<sub>2</sub> species. Expectedly, the presence of an explicit solvent and its contribution toward rearrangement pathway leads to lowering the activation barrier by 160.6 and 109.1 kJ/mol in TS<sub>7b</sub> and TS<sub>8b</sub>, respectively.



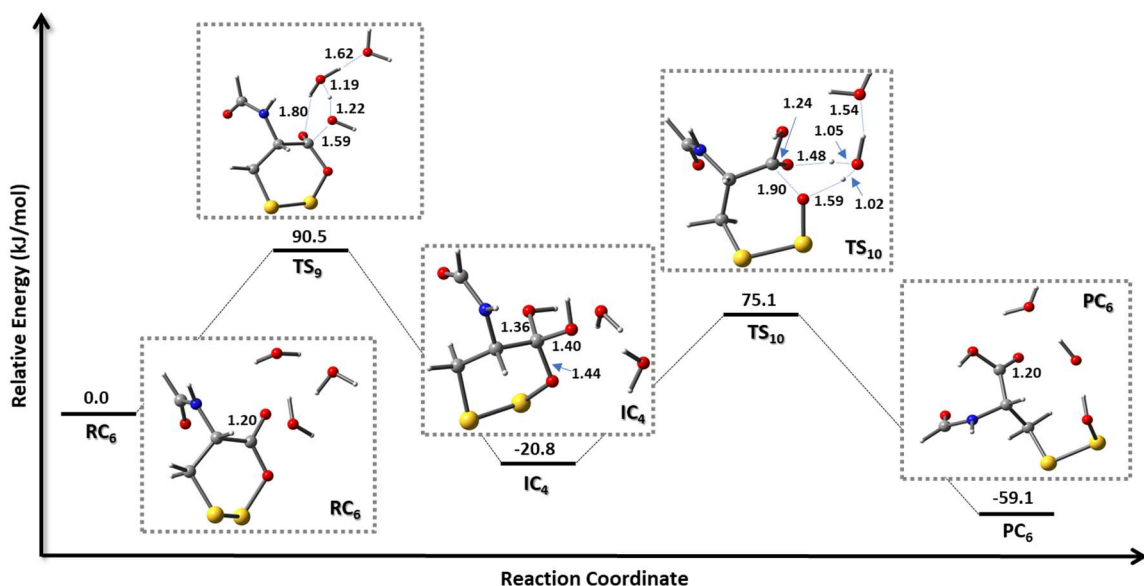
**Figure 6.10.** Potential energy surface with optimized structures and selected bond distances in ångströms (Å) for the formation of Cys-SS<sup>+</sup>(OH)-NH<sub>2</sub>. The black and blue colored surface is for a water-assisted and uncatalyzed reaction.

### 6.3.2.3 Reactions Resulted of Cys-SS-NH<sub>2</sub><sup>+</sup>-OH Derived from C-terminus Cysteiny Persulfide and HNO

Similarly, N-protonated species of Cys-SS-NHOH leads to the formation of a 6-membered cyclic structure and hydroxylamine through a nucleophilic attack of O atom of carboxylate group to the backside of S atom with  $\Delta E = -1138.9$  kJ/mol (**Scheme 6.7**).

**Scheme 6.7.** Schematic representation of Cys-SS-OH formation from Cys-SS-NHOH.

The obtained 6-membered cyclic structure is prone to be hydrolyzed under 2-step water-assisted reaction to yield Cys-SS-OH species (**PC<sub>6</sub>**) which lies 59.1 kJ/mol lower in energy relative to **RC<sub>5</sub>** with two activation barriers of 90.5 (**TS<sub>9</sub>**) and 75.1 (**TS<sub>10</sub>**) kJ/mol and formation of a tetrahedral intermediate (**IC<sub>4</sub>**) (**Figure 6.11**). Comparing the results with cysteine sulfenic acid formation implies kinetic and thermodynamic preference of Cys-SS-OH to cysteine sulfenic acid formation (Cys-S-OH).

**Figure 6.11.** Potential energy surface of the water-catalyzed Cy-SS-OH formation with optimized structures and selected bond lengths in ångstroms (Å).

### 6.3.3 Cys-S-NHOH versus Cys-S-ONH<sub>2</sub>

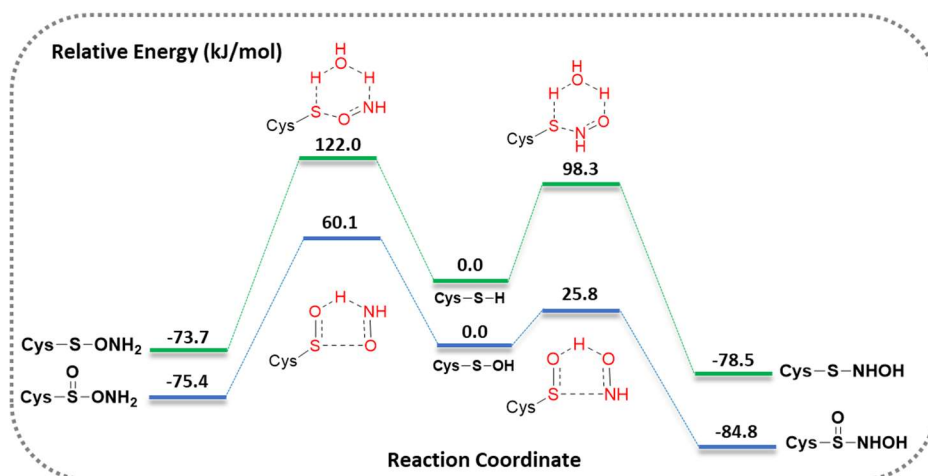
Alternatively, due to higher proton affinity of N atom in HNO relative to O atom, the nucleophilicity of S atom in the internal cysteine residue toward O atom in HNO is also

studied. Solvent catalyzed S—O bond formation and proton transfer from S to N<sub>HNO</sub> atom proceeds *via* a 6-membered cyclic transition structure to form Cys-S-ONH<sub>2</sub> with the  $\Delta E^\ddagger$  and  $\Delta E$  of 122.0 and  $-73.7$  kJ/mol, respectively (see **Figure 6.12**). As results indicate the formation of Cys-S-NHOH should be considered as both kinetic and thermodynamic products compared to Cys-S-ONH<sub>2</sub> formation.

### 6.3.4 Reactivity of HNO toward Cysteine Sulfenic Acid

Reactivity of HNO toward the highly active S atom in cysteine sulfenic acid (Cys-S-OH) is also investigated. As shown in **Figure 6.12**, obtained 5-membered cyclic transition structure resulted from proton transfer of O—H bond in Cys-S-OH to O atom of HNO with activation barrier of 25.8 kJ/mol causes the formation of Cys-S(O)-NHOH species which lies 84.8 kJ/mol lower in energy relative to the reactant complex. Obviously, the activation barrier for the formation of Cys-S(O)-NHOH, comparing Cys-S-NHOH have been decreased dramatically by 72.5 to 25.8 kJ/mol relative to the Cys-NH-OH formation. Furthermore, the reactivity of S atom in cysteine sulfenic acid toward the O atom of HNO is also studied. Expectedly, the activation barrier of formation Cys-S(O)-ONH<sub>2</sub> has been considerably lowered by 61.9 from 122.0 to 60.1 kJ/mol comparing Cys-S-ONH<sub>2</sub> species (**Figure 6.12**). However, the observed changes in the reaction energy are not as noticeable as the activation barrier ( $< 6.3$  kJ/mol).

Observed results clearly reflect the higher chemical tendency of the oxidized S atom in the studied nucleophilic reactions toward O and N atoms in HNO. Although products derived from nucleophilic addition of S to N atom of HNO are both kinetically and thermodynamically favored over nucleophilic addition of S to O atom of HNO.



**Figure 6.12.** Potential energy surface of reactivity of O and N atoms in HNO toward cysteine residue (green colored surface) and cysteine sulfenic acid (blue colored surface) with the schematic representation of transition structures.

**Sulfonamide.H<sup>+</sup> (Cys-S<sup>+</sup>(O)(OH)-NH<sub>2</sub>) Formation:** As well, formation of sulfonamide.H<sup>+</sup> through the arrangement reaction of Cys-S(O)-NH<sub>2</sub><sup>+</sup>OH, a product resulting from the highly observed reactivity of HNO toward cysteine sulfenic acid, was investigated. Solvent-assisted 1,2-H shift *via* either a 3-membered cyclic transition structure ( $\Delta E^\ddagger = 83.2$  kJ/mol) or a 5-membered cyclic transition structure ( $\Delta E^\ddagger = 72.4$  kJ/mol) results in Cys-S<sup>+</sup>(O)=NH formation. The obtained intermediate lies 43.7 kJ/mol lower in energy than the reactant complex. Despite an observed increase in the activation barrier ( $\sim 6$  kJ/mol), the reaction energy for the formation of this species indicates a noticeable decrease of 65.8 kJ/mol with respect to the corresponding reduced form (Cys-S<sup>+</sup>=NH).

Similarly, the obtained species later undergoes a nucleophilic attack by one water molecule and through the formation of a 6-membered cyclic transition structure with a cost of  $-19.8$  kJ/mol relative to reactant complex yields the O-protonated form of sulfonamide (Cys-S<sup>+</sup>(O)(OH)-NH<sub>2</sub>) with  $\Delta E = -117.5$  kJ/mol. It should be noted that despite a significant decrease in the second activation barrier (59.8 kJ/mol) relative to the analogous

reduced form of cysteine residue,  $\text{Cys-S}^+(\text{O})(\text{OH})\text{-NH}_2$  is not considered as the thermodynamic product compared to  $\text{Cys-S}^+(\text{OH})\text{-NH}_2$  formation.

**Thiosulfinate (Cys-S(O)S-Cys) Formation:**  $\text{S}_{\text{N}}2$  nucleophilic substitution of an anionic cysteine ( $\text{Cys-S}^-$ ) toward highly active species of  $\text{Cys-S(O)-NH}_2^+\text{OH}$  leads to formation of thiosulfates ( $\text{Cys-S—S(O)-Cys}$ ) with activation barrier of 62.5 kJ/mol, which in comparison with analogous disulfide ( $\text{Cys-S—S-Cys}$ ) indicates a decrease of 59.5 kJ/mol. In addition, energy reaction decreases from -132 kJ/mol in disulfide formation to -149.8 kJ/mol in thiosulfate formation. In other words, thiosulfate formation ( $\text{S—S(O)}$ ) kinetically and thermodynamically is favored over disulfide ( $\text{S—S}$ ) formation.

## 6.4 Conclusions

Due to the biological significance of HNO-derived post-translational modification of cysteine-containing peptide, the present study has investigated the reactivity of highly reactive HNO toward cysteiny and cysteiny persulfide residues located in the different position in the amino acid sequence.

Our findings show that the formation of  $\text{Cys-X-NHOH}$  ( $\text{X=S}$  and  $\text{S-S}$ ) as the first intermediate in the reaction between the target peptide (cysteiny and cysteiny persulfide) and HNO is independent of the peptide's position. However, the potential energy surfaces for the formation of this species is affected by the peptide's location. As observed, N-terminus and C-terminus cysteiny require the highest and lowest activation barriers for the formation of  $\text{Cys-S-NHOH}$ , respectively, while the reaction energy represents the inverted order.

Furthermore decreasing the polarity of environment from  $\epsilon=78.35$  to 10 and then 4, independent of cysteiny position increases the activation barrier and formation energies. Although the conformational rotation was observed by moving from a polar protein environment ( $\epsilon=10$ ) to a non-polar protein environment ( $\epsilon=4$ ) in the obtained transition



structure for the reaction of N-terminus cysteiny with HNO. Indeed, this rotational change in non-polar protein environment is assumed to facilitate proton transferring process from the target N-terminus cysteiny to HNO which results in a lower activation barrier relative to protein with higher polarity environment.

In addition, comparing the reactivity of cysteiny and cysteiny persulfide toward highly electrophilic HNO represents the higher tendency of Cys-SSH in the reaction with HNO rather than Cys-SH resulted from enhanced nucleophilicity and lower  $pK_a$  of cysteine persulfide.

Our investigations also show the obtained Cys-X-NHOH (X=S and S-S) intermediate, dependence upon the peptide's position (C-terminus or internal peptide) can undergo more modifications. The Cys-X-NH<sub>2</sub><sup>+</sup>OH intermediate, derived from the internal peptide, proceeds through a 2-step rearrangement reaction and forms Cys-X<sup>+</sup>(OH)-NH<sub>2</sub> (X=S and S-S), or a nucleophilic substitution reaction and forms Cys-XX-Cys (X=S).

According to our results, formation of Cys-SS<sup>+</sup>(OH)-NH<sub>2</sub> species despite preferring kinetically, is not thermodynamically favored over Cys-S<sup>+</sup>(OH)-NH<sub>2</sub> formation.

Alternatively, under a nucleophilic attack of excess cysteine residue with a higher activation barrier relative to the rearrangement pathway disulfide form.

On the other hand, Cys-X-NH<sub>2</sub><sup>+</sup>OH (X=S and S-S) derived from the C-terminus residue, leads to formation a 5- or 6-membered cyclic intermediate in cysteiny and cysteiny persulfide, respectively. Indeed, this cyclic intermediate is a result of a nucleophilic attack of O of free carboxylate to S atom. Hydrolysis of these cyclic structures yields Cys-X-OH (X=S and S-S) species. Interestingly, the formation of these new oxidized products take place independent of reactive oxygen species but the peptide's position and intermediacy of mentioned cyclic structures. As a result, the nature of HNO-derived post-translational modification of our target peptides is affected by the position of them in the protein and emphasize the significance of local environment on HNO-derived modification.

In the present study the reactivity of activated S atom in oxidized cysteine, cysteine sulfenic acid, toward electrophilic HNO was also examined. In comparison, the formation of Cys-S(O)-NHOH both thermodynamically and kinetically is preferred to Cys-S-NHOH formation.

## 6.5 References

1. Harmel, R.; Fiedler, D., Features and rRegulation of Non-Enzymatic Post-Translational Modifications. *Nat. Chem. Biol.* **2018**, *14*, 244.
2. Walsh, G., Post-Translational Modifications of Protein Biopharmaceuticals. *Drug Discov. Today*. **2010**, *15*, 773-780.
3. Vlastaridis, P.; Kyriakidou, P.; Chaliotis, A.; Van de Peer, Y.; Oliver, S. G.; Amoutzias, G. D., Estimating the Total Number of Phosphoproteins and Phosphorylation Sites in Eukaryotic Proteomes. *GigaScience* **2017**, *6*, 1-11.
4. Cohen, P., The Origins of Protein Phosphorylation. *Nat. Cell Biol.* **2002**, *4*, 127-130.
5. Carrico, I. S., Chemoselective Modification of Proteins: Hitting the Target. *Chem. Soc. Rev.* **2008**, *37*, 1423-1431.
6. Cruz, E. R.; Nguyen, H.; Nguyen, T.; Wallace, I. S., Functional Analysis Tools for Post-Translational Modification: A Post-Translational Modification Database for Analysis of Proteins and Metabolic Pathways. *Plant J.* **2019**, 1-11.
7. Audagnotto, M.; Dal Peraro, M., Protein Post-Translational Modifications: In Silico Prediction Tools and Molecular Modeling. *Comput. Struct. Biotech. J.* **2017**, *15*, 307-319.
8. Leonard, S. E.; Carroll, K. S., Chemical ‘omics’ Approaches for Understanding Protein Cysteine Oxidation In Biology. *Curr. Opin. Chem. Biol.* **2011**, *15*, 88-102.
9. Reddie, K. G.; Carroll, K. S., Expanding the Functional Diversity of Proteins Through Cysteine oxidation. *Curr. Opin. Chem. Biol.* **2008**, *12*, 746-754.

10. Chung, H. S.; Wang, S. B.; Venkatraman, V.; Murray, C. I.; Van Eyk, J. E., Cysteine Oxidative Posttranslational Modifications Emerging Regulation in the Cardiovascular System. *Circ. Res.* **2013**, *112*, 382-392.
11. Miranda, K. M., The Chemistry of Nitroxyl (HNO) and Implications in Biology. *Coord. Chem. Rev* **2005**, *24*, 433-455.
12. Fukuto, J. M.; Bartberger, M. D.; Dutton, A. S.; Paolocci, N.; Wink, D. A.; Houk, K. N., The Physiological Chemistry and Biological Activity of Nitroxyl (HNO): The Neglected, Misunderstood, and Enigmatic Nitrogen Oxide. *Chem. Res. Toxicol.* **2005**, *18*, 790-801.
13. Flores-Santana, W.; Salmon, D. J.; Donzelli, S.; Switzer, C. H.; Basudhar, D.; Ridnour, L.; Cheng, R.; Glynn, S. A.; Paolocci, N.; Fukuto, J. M.; Miranda, K. M.; Wink, D. A., The Specificity of Nitroxyl Chemistry Is Unique Among Nitrogen Oxides in Biological Systems. *Antioxid. Redox Sign.* **2011**, *14*, 1659-1674.
14. Averill, B. A., Dissimilatory Nitrite and Nitric Oxide Reductases. *Chem. Rev.* **1996**, *96*, 2951-2964.
15. Farmer, P. J.; Sulc, F., Coordination chemistry of the HNO Ligand with Hemes and Synthetic Coordination Complexes. *J. Inorg. Biochem.* **2005**, *99*, 166-184.
16. Doctorovich, F.; Bikiel, D.; Pellegrino, J.; Suárez, S. A.; Larsen, A.; Martí, M. A., Nitroxyl (Azanone) Trapping by Metalloporphyrins. *Coord. Chem. Rev.* **2011**, *255*, 2764-2784.
17. Tocchetti, C. G.; Stanley, B. A.; Murray, C. I.; Sivakumaran, V.; Donzelli, S.; Mancardi, D.; Pagliaro, P.; Gao, W. D.; van Eyk, J.; Kass, D. A.; Wink, D. A.; Paolocci, N., Playing with Cardiac "Redox Switches": The "HNO Way" to Modulate Cardiac Function. *Antioxid. Redox Sign.* **2011**, *14*, 1687-1698.
18. Paolocci, N.; Katori, T.; Champion, H. C.; St John, M. E.; Miranda, K. M.; Fukuto, J. M.; Wink, D. A.; Kass, D. A., Positive Inotropic and lusitropic Effects of HNO/NO- in

Failing Hearts: Independence from beta-Adrenergic Signaling. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5537-5542.

19. Sabbah, H. N.; Tocchetti, C. G.; Wang, M. J.; Daya, S.; Gupta, R. C.; Tunin, R. S.; Mazhari, R.; Takimoto, E.; Paolocci, N.; Cowart, D.; Colucci, W. S.; Kass, D. A., Nitroxyl (HNO). *Circ-Heart Fail.* **2013**, *6*, 1250-1258.

20. Sidorkina, O.; Espey, M. G.; Miranda, K. M.; Wink, D. A.; Laval, J., Inhibition of Poly(ADP-RIBOSE) Polymerase (PARP) by Nitric Oxide and Reactive Nitrogen Oxide Species. *Free Radic. Biol. Med.* **2003**, *35*, 1431-1438.

21. Kim, W. K.; Choi, Y. B.; Rayudu, P. V.; Das, P.; Asaad, W.; Arnelle, D. R.; Stamler, J. S.; Lipton, S. A., Attenuation of NMDA Receptor Activity and Neurotoxicity by Nitroxyl Anion, NO. *Neuron* **1999**, *24*, 461-469.

22. Miranda, K. M.; Paolocci, N.; Katori, T.; Thomas, D. D.; Ford, E.; Bartberger, M. D.; Espey, M. G.; Kass, D. A.; Feelisch, M.; Fukuto, J. M.; Wink, D. A., A Biochemical Rationale for the discrete Behavior of Nitroxyl and Nitric Oxide in the Cardiovascular System. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9196-9201.

23. DeMaster, E. G.; Redfern, B.; Nagasawa, H. T., Mechanisms of Inhibition of Aldehyde Dehydrogenase by Nitroxyl, the Active Metabolite of the Alcohol Deterrent Agent Cyanamide. *Biochem. Pharmacol.* **1998**, *55*, 2007-2015.

24. Wong, P. S. Y.; Hyun, J.; Fukuto, J. M.; Shiota, F. N.; DeMaster, E. G.; Shoeman, D. W.; Nagasawa, H. T., Reaction between S-Nitrosothiols and Thiols: Generation of Nitroxyl (HNO) and Subsequent Chemistry. *Biochemistry* **1998**, *37*, 5362-5371.

25. Doyle, M. P.; Mahapatro, S. N.; Broene, R. D.; Guy, J. K., Oxidation and Reduction of Hemoproteins by Trioxodinitrate(II). The Role of Nitrosyl Hydride and Nitrite. *J. Am. Chem. Soc.* **1988**, *110*, 593-599.

26. Shen, B.; English, A. M., Mass spectrometric Analysis of Nitroxyl-Mediated Protein Modification: Comparison of Products Formed with Free and protein-Based Cysteines. *Biochemistry* **2005**, *44*, 14030-14044.

27. Keceli, G.; Toscano, J. P., Reactivity of Nitroxyl-Derived Sulfinamides. *Biochemistry* **2012**, *51*, 4206-4216.
28. Jackson, M. I.; Fields, H. F.; Lujan, T. S.; Cantrell, M. M.; Lin, J.; Fukuto, J. M., The Effects of Nitroxyl (HNO) on H<sub>2</sub>O<sub>2</sub> Metabolism and Possible Mechanisms of HNO signaling. *Arch. Biochem. Biophys.* **2013**, *538*, 120-129.
29. Filipovic, M., 6 - HNO–Thiol Relationship. In *The Chemistry and Biology of Nitroxyl (HNO)*, Doctorovich, F.; Farmer, P. J.; Marti, M. A., Eds. Elsevier: Boston, 2017; pp 105-126.
30. Vaananen, A. J.; Kankuri, E.; Rauhala, P., Nitric Oxide-Related Species-Induced Protein Oxidation: Reversible, Irreversible, and Protective Effects on Enzyme Function of Papain. *Free Radic. Biol. Med.* **2005**, *38*, 1102-1111.
31. Lopez, B. E.; Rodriguez, C. E.; Pribadi, M.; Cook, N. M.; Shinyashiki, M.; Fukuto, J. M., Inhibition of yeast Glycolysis by Nitroxyl (HNO): A Mechanism of HNO Toxicity and Implications to HNO biology. *Arch. Biochem. Biophys.* **2005**, *442*, 140-148.
32. Lopez, B. E.; Wink, D. A.; Fukuto, J. M., The Inhibition of Glyceraldehyde-3-phosphate Dehydrogenase by Nitroxyl (HNO). *Arch. Biochem. Biophys.* **2007**, *465*, 430-436.
33. Cook, N. M.; Shinyashiki, M.; Jackson, M. I.; Leal, F. A.; Fukuto, J. M., Nitroxyl-Mediated Disruption of Thiol Proteins: Inhibition of the Yeast Transcription Factor Ace1. *Arch. Biochem. Biophys.* **2003**, *410*, 89-95.
34. Fukuto, J. M.; Bianco, C. L.; Chavez, T. A., Nitroxyl (HNO) signaling. *Free Radic. Biol. Med.* **2009**, *47*, 1318-1324.
35. Fukuto, J. M.; Carrington, S. J., HNO Signaling Mechanisms. *Antioxid. Redox Sign.* **2011**, *14*, 1649-1657.
36. Suarez, S. A.; Muñoz, M.; Alvarez, L.; Venâncio, M. F.; Rocha, W. R.; Bikiel, D. E.; Marti, M. A.; Doctorovich, F., HNO Is Produced by the Reaction of NO with Thiols. *J. Am. Chem. Soc.* **2017**, *139*, 14483-14487.

37. Flores-Santana, W.; Switzer, C.; Ridnour, L. A.; Basudhar, D.; Mancardi, D.; Donzelli, S.; Thomas, D. D.; Miranda, K. M.; Fukuto, J. M.; Wink, D. A., Comparing the Chemical Biology of NO and HNO. *Arch. Pharm. Res.* **2009**, *32*, 1139-1153.
38. Keceli, G.; Toscano, J. P., Reactivity of C-Terminal Cysteines with HNO. *Biochem.* **2014**, *53*, 3689-3698.
39. Zhang, Y., Computational Investigations of HNO in Biology. *J. Inorg. Biochem.* **2013**, *118*, 191-200.
40. Zhou, C.; Liang, J.; Cheng, S.; Shi, T.; Houk, K. N.; Wei, D.-Q.; Zhao, Y.-L., Ab Initio Molecular Metadynamics Simulation for S-nitrosylation by Nitric Oxide: S-Nitroxide as the Key Intermediate. *Mol. Simulat.* **2017**, *43*, 1134-1141.
41. Sherman, M. P.; Grither, W. R.; McCulla, R. D., Computational Investigation of the Reaction Mechanisms of Nitroxyl and Thiols. *J. Org. Chem.* **2010**, *75*, 4014-4024.
42. Ivanova, L. V.; Cibich, D.; Deye, G.; Talipov, M. R.; Timerghazin, Q. K., Modeling of S-Nitrosothiol-Thiol Reactions of Biological Significance: HNO Production by S-Thiolation Requires a Proton Shuttle and Stabilization of Polar Intermediates. *Chembiochem* **2017**, *18*, 726-738.
43. Park, C. M.; Weerasinghe, L.; Day, J. J.; Fukuto, J. M.; Xian, M., Persulfides: Current Knowledge and Challenges in Chemistry and Chemical Biology. *Mol. Biosyst.* **2015**, *11*, 1775-1785.
44. Iciek, M.; Kowalczyk-Pachel, D.; Bilska-Wilkosz, A.; Kwiecien, I.; Gorny, M.; Wlodek, L., S-sulphydration as a Cellular Redox Regulation. *Biosci. Rep.* **2016**, *36*, e00304.
45. Yadav, P. K.; Martinov, M.; Vitvitsky, V.; Seravalli, J.; Wedmann, R.; Filipovic, M. R.; Banerjee, R., Biosynthesis and Reactivity of Cysteine Persulfides in Signaling. *J. Am. Chem. Soc.* **2016**, *138*, 289-299.
46. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.;

Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; N. Rega; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; J. V. Ortiz; Cioslowski, J.; Fox, D. J. *Gaussian 09*, E.01; Gaussian, Inc.: Wallingford CT, 2015.

47. Zhao, Y.; Truhlar, D. G., The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215-241.

48. Nikoo, S.; Meister, P. J.; Hayward, J. J.; Gault, J. W., An Assessment of Computational Methods for Calculating Accurate Structures and Energies of Bio-Relevant Polysulfur/Selenium-Containing Compounds. *Molecules* **2018**, *23*, 3323.

49. Siegbahn, P. E. M.; Himo, F., The Quantum Chemical Cluster Approach for Modeling Enzyme Reactions. *WIREs. Comput. Mol. Sci.* **2011**, *1*, 323-336.

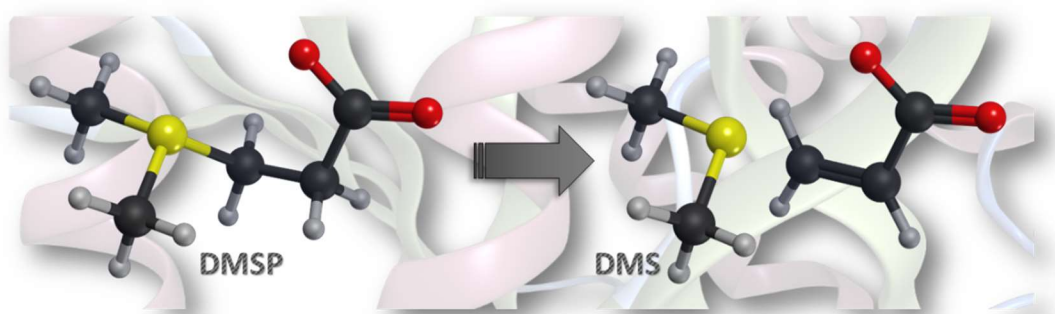
50. AIMAll (Version 17.11.14); Todd A. Keith, TK Gristmill Software: Overland Park KS, USA, 2017.

51. Liu, H.; Gault, J. W., Protonation of Guanine Quartets and Quartet Stacks: Insights from DFT Studies. *Phys. Chem. Chem. Phys.* **2009**, *11*, 278-287.

52. Popelier, P. L. A., Characterization of a Dihydrogen Bond on the Basis of the Electron Density. *J. Phys. Chem. A* **1998**, *102*, 1873-1878.

## CHAPTER 7

# Computational Study on the Mechanism of Bacterial Ni(II)-Dimethylsulfoniopropionate Lyase





## 7.1 Introduction

Dimethylsulfoniopropionate (DMSP) is one of the most globally ubiquitous sulfur-containing compounds.<sup>1-3</sup> It is biosynthesized by a range of marine organisms including phytoplankton, algae, and corals, as well as some plants. Within such organisms, in particular, marine organisms, it is thought to act as an osmolyte, antioxidant, and predator deterrent.<sup>4-9</sup> In addition, however, it can be used as a sulfur and carbon source in some marine organisms being catabolized through either demethylation or cleavage pathways.<sup>8, 10-12</sup> The former, degradation *via* demethylation, is the most common pathway due in part to, for instance, the high abundance of SAR11 bacteria (a plankton).<sup>13</sup> It leads to a range of products including methylthiol, methyl mercaptopropionate, and acetaldehyde.

In contrast, in the cleavage pathway DMSP is catalytically degraded by DMSP lyases to dimethyl sulfide (DMS) and acrylate. Despite the dominance of alternate demethylation pathway, the DMSP cleavage (DMS formation) pathway is of considerable importance both biologically and globally. Indeed, DMS is the largest natural source of atmospheric sulfur, comparable in quantity to the SO<sub>2</sub> formed by coal combustion.<sup>14-15</sup> For example, DMS has been found to act as a signaling molecule in some marine organisms and seabirds.<sup>16-19</sup> Meanwhile, DMS plays significant roles in global sulfur and carbon cycling as well as climate regulation *via* its involvement in cloud formation.<sup>11, 20-21</sup> It is also noted that DMS and acrylate are also industrially important with a broad range of applications.<sup>22</sup>

To date, several distinct functional bacterial DMSP lyase metalloenzymes have been identified from proteobacteria, fungi, and phytoplankton, including DddD, DddY, DddP, DddL, DddQ, and DddW, and Alma1 from algae and corals.<sup>8, 11, 21, 23-30</sup> In each, the metal ion(s) are thought to be important for the enzymes function. Notably, they exhibit a range of metal ion preference both in valency and number of metal ions in the active site: Fe(II) or Mn(II) in DddW;<sup>10</sup> Zn(II) or Fe(III) in DddQ;<sup>23, 25</sup> and Fe(II) in the binuclear iron DddP.<sup>30-31</sup> Recently, a metalloenzyme DMSP lyase DddK was discovered in SAR11 and similarly exhibits promiscuous metal ion binding with a preference for divalent transition

metal ions.<sup>32</sup> Experimentally, for instance, it has been shown to be able to bind Ni(II), Mn(II), Fe(II), Co(II), Zn(II), Cu(II), and Fe(III) within its active site.<sup>32</sup> The catalytic rate activity of DddK is maximal when Ni(II) is the metal ion ligated in the active site.<sup>32</sup> It is noted that Ni(II) is generally one of the most plentiful metal ions in ocean waters.<sup>33-34</sup>

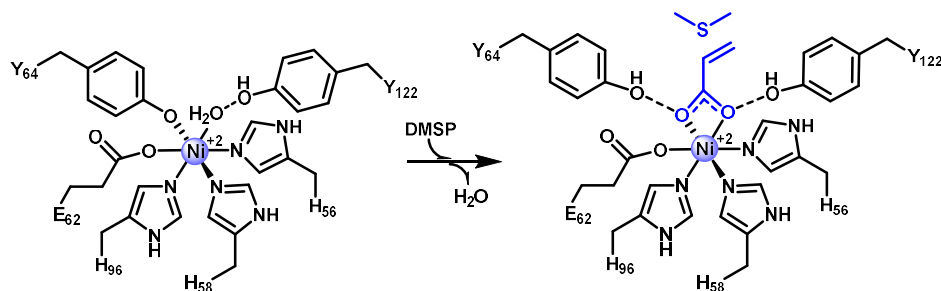
To date, experimental X-ray crystal structures of two different forms of DddK have been determined: a mononuclear Ni-DddK and a dinuclear FeZn-DddK.<sup>32</sup> Consequently, they exhibit different metal ion active site-coordinations, and DMSP lyase activities.<sup>32</sup> For example, in FeZn-DddK (PDB ID: 5TG0), similar to that observed in Fe(III)-DddQ and Zn(II)-DddQ,<sup>23, 25</sup> the Fe(III) coordinates monodentately with each of two histidyls and a tyrosyl, and bidentately with a glutamyl, and one water molecule. Meanwhile, the Zn(II) coordinates monodentately with three histidyls and a glutamyl residue, and one water molecule. As a result, the Fe(III) and Zn(II) have octahedral and trigonal bipyramidal geometries, respectively.

In contrast, in the absence of substrate, the metal ion in Ni-DddK (PDB ID: 5TFZ) coordinates in a manner similar to that of the Zn(II) in FeZn-DddK; monodentately with three histidyl residues (His56, 58, and 96) and a glutamyl (Glu62), and a water molecule. However, it additionally ligates to a tyrosyl (Tyr64) *via* its deprotonated R-group hydroxyl oxygen ( $\text{Tyr}_{64}\text{O}^-$ ). As a result, the active site Ni(II) has octahedral coordination. Based on X-ray crystal structures it has been proposed that the substrate, DMSP, binds bidentately to the Ni(II) center, displacing both the metal ion-ligated  $\text{H}_2\text{O}$  and Tyr64. Furthermore, the now displaced Tyr64 and Tyr122 contribute to substrate binding through interactions between their R-group oxygen's and the substrates carboxyl (*i.e.*,  $\text{subO} \cdots \text{O}_{\text{Tyr64}}$  and  $\text{subO} \cdots \text{O}_{\text{Tyr122}}$ ). The  $\text{Tyr}_{64}\text{O}^-$  is then able to function as a Lewis base and abstract a proton from the substrate's  $-\text{C}_\alpha\text{H}_2-$  group, thus breaking  $\text{C}_\beta-\text{S}$  bond *via* a  $\beta$ -elimination mechanism, and resulting in formation of DMS and acrylate (**Scheme 7.1**).

There have been few computational investigations on the mechanism(s) of DMSP lyase. Recently, Wang *et al.* examined the mechanism of a binuclear iron DddP, R/DddP.<sup>31</sup> More

specifically, they employed a quantum mechanics (QM)-chemical cluster approach in combination with the hybrid DFT method B3LYP, corrected for dispersion, to examine catalytic mechanisms for DMSP decomposition by the homodimeric metalloenzyme. They concluded that an aspartyl acts as a base to abstract a  $-C_{\alpha}H_2-$  proton, and that the mechanism preferably proceeds *via* a concerted  $\beta$ -elimination with a barrier of  $\sim 17$  kcal/mol. To date, however, to the best of our knowledge, no computational studies have examined the mechanism of a mononuclear DMSP lyases. Thus, unfortunately, despite their industrial, environmental, and biological importance a number of key questions remain unclear about these important enzymes, in particular, the highly promiscuous metal binding DddK enzymes.

**Scheme 7.1.** Schematic representation of the proposed Ni-DddK catalysed decomposition of DMSP.<sup>32</sup>



In this present study, we have used a multi-scale computational approach to investigate the mechanism of Ni-DddK. More specifically, molecular dynamics simulation (MD), QM-cluster and hybrid quantum mechanics/molecular mechanics (QM/MM) have been applied to obtain greater insights into substrate binding, as well as the enzyme mechanism, including the roles of the environment and key active site residues. In addition, the effects of substituting Ni(II) with Mn(II) on the catalytic mechanism has been examined.

## 7.2 Computational Methods

### 7.2.1 Molecular Dynamics (MD) Simulations

The Molecular operating environment (MOE) program<sup>35</sup> was used to prepare the chemical model for MD simulations. The experimental X-ray crystal structure of a Ni(II)-containing DddK from SAR11 complexed with diacrylate (PDB ID: 5TFZ<sup>32</sup>) was used as the initial template. Protonation states of all residues were assigned based on the PROPKA protonation tool available in MOE. Two different chemical models were prepared: the unbound active site (obtained by omitting the diacrylate) and the bound active site (obtained by mutation of diacrylate to give DMSP).

Each enzyme complex was then solvated and surrounded by a water layer 6 Å deep. The resulting solvated models were then minimized using the AMBER14:EHT molecular mechanics (MM) forcefield. To help ensure the experimentally observed octahedral coordination of Ni(II) was maintained during the MD simulations, the ligated residues to Ni(II) were kept fixed. That is, in the unbound complex His56, His58, His96, Glu62, Tyr64, and the Ni(II)-bound water were held fixed. Meanwhile, in the substrate-bound complex the His56, His58, His96, and Glu62 residues were held fixed, as was the bidentately Ni(II)-ligated substrate (DMSP). All MD simulations were performed using the NAMD program.<sup>36</sup> The final minimized systems were then submitted for 100 ps equilibrium MD simulations in which the temperature as gradually raised from 0 to 300 K at constant pressure. These were then followed by 1 ns production MD simulations with a time step of 2 fs under constant pressure and temperature. The conformations generated for each complex from the 1 ns MD simulations were then analyzed and clustered according to their root mean square deviation (RMSD). A representative structure for each complex was then chosen from the most populated cluster to provide a suitable initial chemical model for the subsequent QM-cluster and QM/MM calculations (see 7.2.2).

### 7.2.2 QM-Cluster and QM/MM Calculations

The Gaussian09 software package<sup>37</sup> was used for all the QM-cluster and ONIOM(QM/MM) calculations. The QM-cluster method is an established approach for obtaining insights into enzymatic mechanisms, in particular those of metalloenzymes.<sup>38</sup> In addition, in a multi-scale approach it is able to help provide additional insights through comparison with results obtained from, for example, QM/MM. Hence, we first used the QM-cluster method to examine the mechanism of DMSP lyase. The chemical clusters used for both the substrate-bound and unbound active sites included the side chains of all residues ligated to the Ni(II) center in the unbound complex. That is, they included His56, His58, His96, Glu62, and Tyr64. The side-chain of Tyr122 was also included as was the Ni(II)-ligated water in the unbound active site (see **Figure 7.1**). To maintain the structural integrity of the truncated QM-cluster model the Ca center was held fixed at its MD-derived position and capped with hydrogens.

The mechanism was then examined using above chemical cluster models. Optimized geometries were obtained using the two *meta*-GGA functionals M06 (27% HF) and M06L (0% HF),<sup>39</sup> in combination with the systematic series of basis sets: 6-31G(d), 6-31G(d,p), and 6-311G(d,p). This enabled us to assess their abilities to reliably and accurately describe the mechanism and consequently provide further support for the results obtained. Harmonic vibrational frequencies of each optimized complex were also determined at the same level of theory. These were used to ensure and characterize each structure as an energy minimum or transition structure and to obtain the Gibbs free energy correction ( $\Delta G_{\text{corr}}$ ) for each structure. Relative Gibbs Free energies for each optimized structure were obtained using the above DFT methods in combination with the 6-311+G(2df,p) basis set, and corrected using the appropriate  $\Delta G_{\text{corr}}$ . It is noted that the polar protein environment was included in all calculations by use of the IEFPCM solvation model with the commonly used value for the dielectric constant ( $\epsilon$ ) of 4.0.<sup>38</sup>

Subsequently, the two-layer ONIOM(QM/MM) method, as implemented in Gaussian09, was used to further examine the mechanism of DMSP lyase and the impact of the protein environment.<sup>40-41</sup> The reactive (QM) regions again included the key residues in and near the active site and were similar to the corresponding QM-cluster chemical models. This region was described using the M06 and M06L<sup>39</sup> methods, in combination with the same range of basis sets as for the QM-cluster studies. The remainder of the protein, the MM region, was treated using the AMBER96 forcefield<sup>42</sup> within the mechanical embedding formalism (ME). It should be noted that to ensure the conformational integrity of the chosen model, the  $\alpha$ -carbon atoms of all residues in the MM layer were kept fixed at their initial optimized position. Harmonic vibrational frequencies and Gibbs free energy corrections were again obtained at the optimization level of theory. Relative Gibbs free energies were obtained by performing single point energy calculations at the ONIOM(DFT/6-311+G(2df,p);AMBER96) on the above optimized structures with inclusion of the appropriate  $\Delta G_{\text{corr}}$ , within both the mechanical (ME) and electrostatic embedding (EE) formalisms.

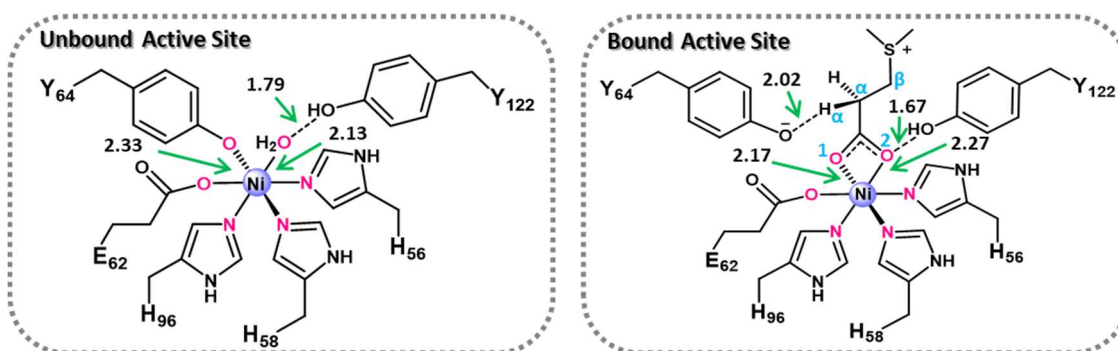
### 7.3 Results and Discussion

#### 7.3.1 Ni-DddK Enzyme

##### 7.3.1.1 QM-Cluster Investigation

For the mechanistic and geometrical investigation of DMSP decomposition, two generated models of unbound and bound active site from MD simulations were optimized using the M06/6-31G(d,p), SCRF=(IEFPCM,eps=4) level of theory. It is noted that optimized structures were obtained for the Ni(II)-containing unbound and bound active sites in both the open-shell singlet and triplet states. For both active site models, the triplet state was found to be lower in energy. Hence, only results obtained for the triplet state are discussed herein.

As shown in **Figure 7.1**, in the unbound form, Ni(II) is ligated to His56, 58, and 96, Glu62 in a monodentate fashion and deprotonated Tyr64. The remaining open coordination is occupied with one molecule of water leading to a 6-coordinated metal center with octahedral geometry. All coordinated ligands are 2.02-2.33 Å from the nickel center. While in the presence of DMSP in the active site, the substrate coordinates to the metal ion *via* its carboxylate end in a bidentate fashion to keep the octahedral geometry and Tyr64 and water molecule have no longer coordination to Ni(II). All ligands coordination distance to the metal center in the active site model is 1.96-2.27 Å.



**Figure 7.1.** Schematic illustration of the optimized structures of the unbound and bound active sites using the QM-cluster approach with the selected bond distances in ångstroms (Å).

As observed, the most significant structural changes of the unbound and bound forms are related to substrate binding in the active model in which  $O_{Y64}$  and  $O_{\text{water}}$  in the pre-active form are replaced with two oxygen atoms of the substrate while other coordination distances with Ni(II) center have not changed considerably.

The “Tyr-switching” mechanism and the detachment of metal-coordinated Tyr were previously investigated as a key initial step in the activity of the astacins family simultaneously upon substrate binding.<sup>43</sup> In a similar fashion, the presence of the DMSP results in the removal of deprotonated Tyr64 from the first metal coordination sphere and

moving it to the second sphere of metal coordination in the enzyme. Also, Tyr64 separation may accommodate steric crowding on substrate binding.

In this rearrangement, the tyrosinate oxygen atom of Tyr64 is positioned at 2.02 Å from  $H_\alpha$  of DMSP and  $H_{Y122}$  is placed in a 1.67 Å distance from the second oxygen atom of the substrate (see **Figure 7.1**). In other words, both Tyr122 and Tyr64 through a strong hydrogen binding stabilize the enzyme-substrate complex in the enzyme active site.

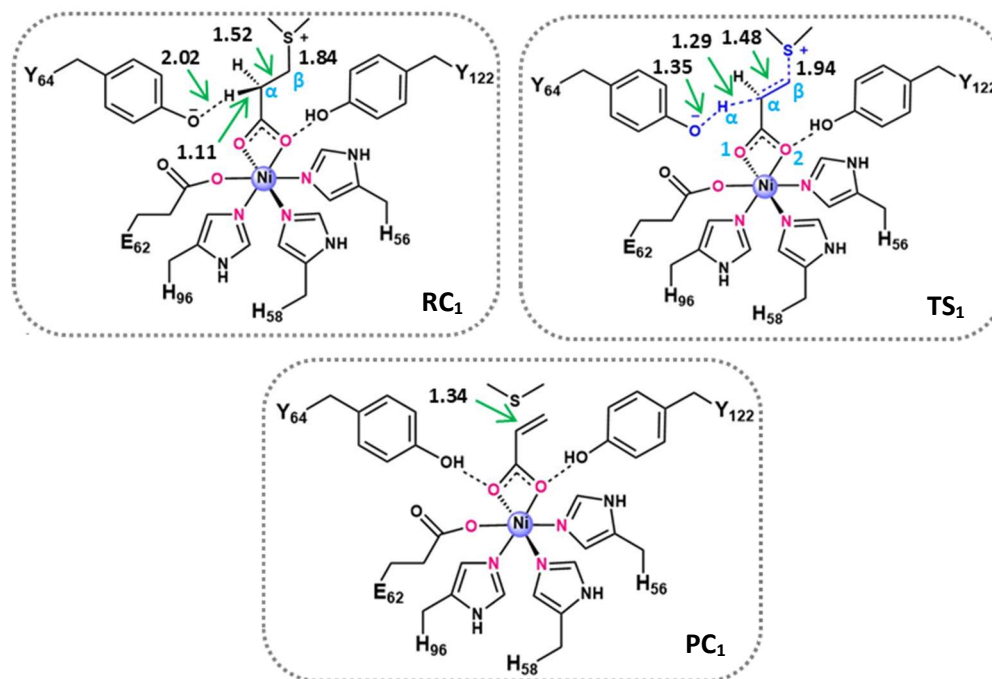
The  $pK_a$  values of both Tyr64 and 122 in the resting and active states of the enzyme have been calculated using the PROPKA 3.0 program.<sup>44</sup> It has been observed that the presence of substrate in the active site lowers the  $pK_a$  of Tyr64 and 122 relative to the enzyme with the unbound active site. Also, it is worth mentioning that the  $pK_a$  of Tyr122 is slightly more affected by the presence of DMSP which makes it act as a stronger H-bonding donor to contribute to substrate binding. In addition, higher observed  $pK_a$  for both Tyr64 and 122 comparing other Tyr residues in the enzyme, may refer to their strong tendency toward abstracting the proton and being neutral. On the other hand, well-positioning of deprotonated Tyr64 in the active form, makes it directly behaves as a Lewis base to abstract a proton of the  $C_\alpha$  atom of DMSP to initiate a  $\beta$ -elimination reaction to produce DMS and acrylate.

Geometrical analysis of the obtained transition structure indicates that the considerable structural changes are related to bonds involving breaking and forming:  $Y64O \cdots H_\alpha$ ,  $C_\alpha \cdots H_\alpha$ ,  $C_\alpha \cdots C_\beta$  and  $S \cdots C_\beta$  relative to the active form, not Ni(II)-centered coordination (coordination distances in **TS1**: 1.96-2.24 Å). As shown in **Figure 7.2**,  $H_\alpha$  atom abstraction by the tyrosinate oxygen of Tyr64 and  $C_\beta$ —S bond cleavage is followed by a concerted single-step mechanism and subsequently releasing dimethyl sulfide and acrylate. The product complex also displays no considerable changes in Ni(II) coordinated ligands (coordination distances in **PC1**: 1.97-2.22 Å).

The obtained free energy surface for the DMSP lyase is also shown in **Figure 7.3A**. The influence of improving the basis set from 6-31G(d) to 6-31G(d,p) and then to 6-311G(d,p)



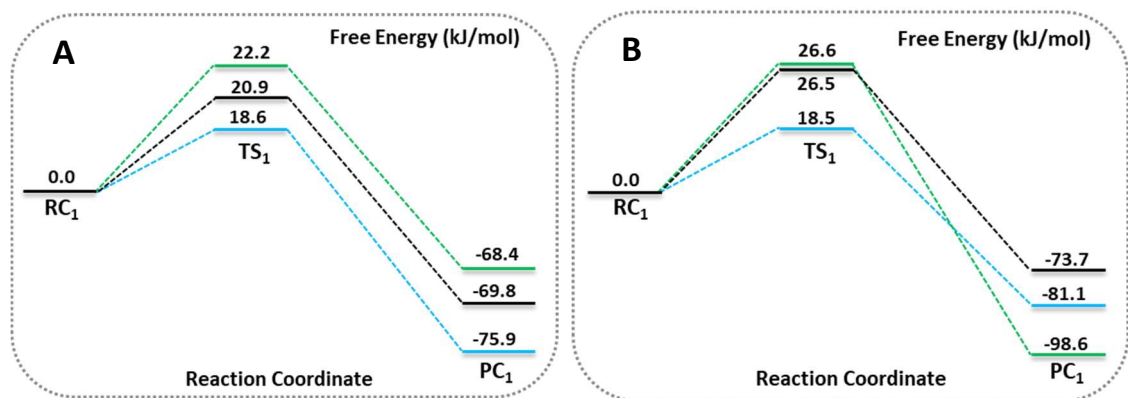
has been studied. As anticipated, by increasing the basis set, free activation energy and free energy of reaction are decreased by 3.6 from 22.2 to 18.6 kJ/mol and 7.5 from  $-68.4$  to  $-75.9$  kJ/mol, respectively. The low activation barrier also reflects that transition structure occurs early along with this reaction, as we observe in the **TS<sub>1</sub>** that  $C_\alpha-C_\beta$  distance has been shortened only by 0.04 Å and the  $r(C_\beta-S)$  has been lengthened by 0.10 Å.



**Figure 7.2.** Schematic illustration of the optimized structures of the reactive complex, transition structure, and product complex using the QM-cluster approach with selected bond distances in ångstroms (Å).

To assess the ability of the M06L functional with 0% HF contribution, it has also been employed to study this concerted  $\beta$ -elimination reaction (see **Figure 7.3B**). A close agreement between the relative free energy obtained using the M06 and M06L functionals is observed, specifically applying the largest basis set (6-311G(d,p)). Also, using M06L functional the free energy of activation is in a good agreement with the M06 functional with a variance of  $\leq 5.6$  kJ/mol. However, it lowers the free energy of the reaction by 30.2 kJ/mol employing the smallest size of basis set, 6-31G(d), with respect to the corresponding

obtained value using M06 functional. It should be noted that the geometrical parameters using these two methods are in a good agreement with a variance of  $< 0.05 \text{ \AA}$ .



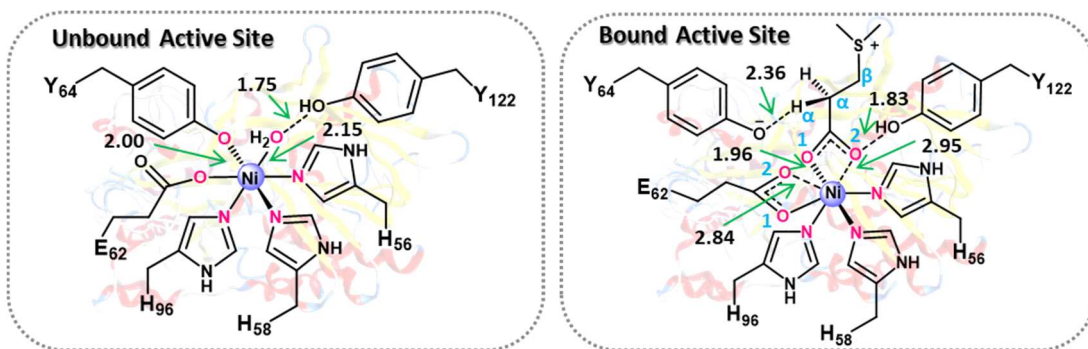
**Figure 7.3.** Free energy surface obtained using the QM-cluster approach; (A) M06 functional (B) M06L functionals. The green, black, and blue colored surfaces are for the 6-31G(d), 6-31G(d,p), and 6-311G(d,p) level of optimization, respectively.

#### 7.3.1.2 QM/MM Investigation

QM/MM optimized structures of the unbound and bound active site models at the ONIOM(M06/6-31G(d,p);AMBER96) level of theory with selected distances in ångstroms (Å) are shown in **Figure 7.4**. The main difference between the two models is attributed to the Ni(II) geometry and coordination. In the unbound active site, His56, 58, and 96, Glu62 in a monodentate fashion and deprotonated Tyr64 with one molecule of water are ligated to Ni(II) that make the metal form an octahedral geometry in the resting state. Upon addition of DMSP to the active site, the metal-bound water and Tyr64 are displaced:  $r_{(\text{Y64O} \cdots \text{Ni})}$ : 4.86 Å) and DMSP coordinates in a monodentate fashion by one oxygen atom of its end carboxylate group:  $r_{(\text{SubO}^1 \cdots \text{Ni})}$ : 1.96 Å vs.  $r_{(\text{SubO}^2 \cdots \text{Ni})}$  is 2.95 Å, unlike the obtained DMSP coordination in the QM-cluster study, in which two oxygen atoms of DMSP were ligated to Ni(II). The second carboxylate oxygen is stabilized by Tyr122 as an H-bonding donor which is positioned at 1.83 Å from  $\text{O}^2_{\text{Sub}}$ . This DMSP monodentate binding fashion *via* one carboxylate oxygen was also observed in DMSP-bound Fe(III)-

DddQ enzyme.<sup>25</sup> Moreover, the deprotonated Tyr64, which presence of substrate has located it in the catalytic state of enzyme, assists in the substrate binding through an H-bonding with  $H_\alpha$  of the substrate with a distance of 2.36 Å. The hydrophobic resting of DMSP is placed in a rather hydrophobic pocket mostly shaped by Leu53, Trp26, and 110. All mentioned interactions help the position of DMSP in the enzyme active site.

It should be noted that the  $E_{62}O^2 \cdots Ni$  distance in the active form has decreased by 0.43 Å from 3.27 Å in the resting state to 2.84 Å (see **Figure 7.4**) to form a distorted octahedral Ni-centered geometry while other Ni-coordinated ligands are at a distance of 1.96-2.20 Å from the metal center.



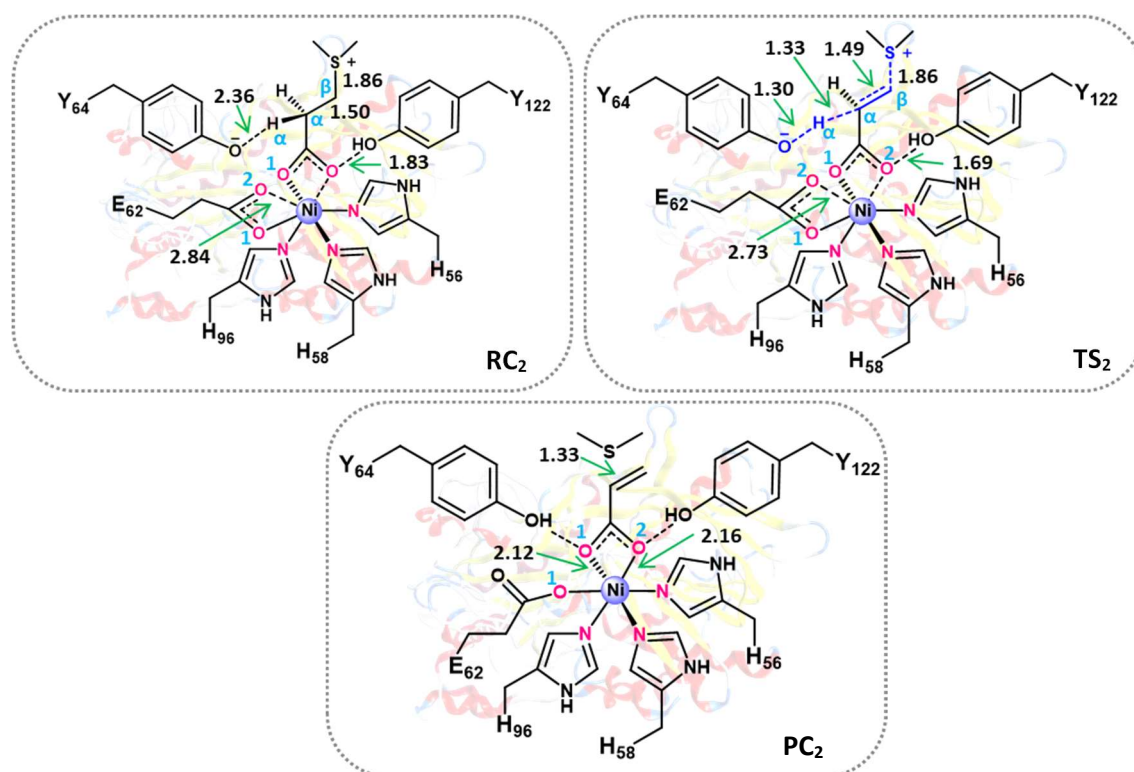
**Figure 7.4.** Schematic representation of the optimized unbound and bound active site complexes using the ONIOM(QM/MM) approach, with selected bond distances in ångstroms (Å).

Besides an anionic Tyr64 role in substrate binding, geometrical studies reveal that proper positioning of this residue at a distance of 3.11 Å from  $C_\alpha$  of DMSP, makes it the most likely candidate to act as a Lewis base to initiate the elimination reaction for abstracting the  $H_\alpha$  from  $C_\alpha$  methylene atom of substrate and releasing DMS and acrylate. In addition, observing no water molecule in the bound active site of the enzyme shows that Tyr64 directly behaves as a Lewis base for abstracting the  $H_\alpha$  (see **Figure 7.5**).

The DMSP cleavage mechanism proceeds through a concerted transition structure at a cost of 28.4 kJ/mol relative to **RC**<sub>2</sub>. As shown in **Figure 7.5**, more specifically, Tyr64 has

deprotonated the  $C_\alpha$  methylene of the substrate such that the  $Y_{64}O \cdots H_\alpha$  is now 1.30 Å. Concomitantly,  $C_\alpha-C_\beta$  has only been slightly shortened by 0.10 to 1.49 Å which with the low computed free activation energy reflects an early transition structure along this reaction. It is worth mentioning that none of our computational investigations prove a theory of stepwise mechanism for DMSP decomposition.

In **TS<sub>2</sub>**, the Tyr122 and  $O^2_{\text{Sub}}$  distance is 1.69 Å which relative to **RC<sub>2</sub>** represents an increased tendency of Tyr122 to stabilize DMSP in the transition structure. In addition,  $r(\text{Ni} \cdots \text{N}_{\text{H96,58,56}})$ ,  $r(\text{Ni} \cdots \text{O}^1_{\text{Sub}})$ , and  $r(\text{Ni} \cdots \text{O}^1_{\text{E62}})$  are not changed significantly compared to **RC<sub>2</sub>**. However  ${}_{\text{E62}}\text{O}^2 \cdots \text{Ni}$  distance is shortened by 0.11 Å to 2.73 Å in **TS<sub>2</sub>**.



**Figure 7.5.** Schematic drawing of the optimized molecular optimized structures of the reactive complex, transition structure, and product complex using the ONIOM(QM/MM) approach with selected bond distances in ångstroms (Å).

This transition structure leads to the formation of DMS and acrylate (**PC<sub>2</sub>**) that lies 240.6 kJ/mol lower in energy than **RC<sub>2</sub>** (see **Figure 7.5**). The latter, in the active site, is ligated

to Ni(II) *via* two carboxylate oxygen atoms with a distance of  $r(\text{Ni}\cdots\text{O}^1_{\text{Acr}})$ : 2.12 Å and  $r(\text{Ni}\cdots\text{O}^2_{\text{Acr}})$ : 2.16 Å. Although in the Fe-DddQ enzyme has been observed that the acrylate ligates the Fe(III) center in a monodentate fashion.<sup>25</sup> Two Tyr64 and 122 also significantly contribute to the acrylate placement in the active site by strong H-binding:  $(\text{AcrO}^1\cdots\text{H}_{\text{Y64}})$ : 1.79 Å and  $(\text{AcrO}^2\cdots\text{H}_{\text{Y122}})$ : 1.69 Å. The remaining Ni(II) coordination is occupied by His56, 58, and 96, and Glu62 in a monodentate fashion with a distance of 2.02-2.18 Å from metal center, resulting in a six coordinated-metal site with the octahedral geometry.

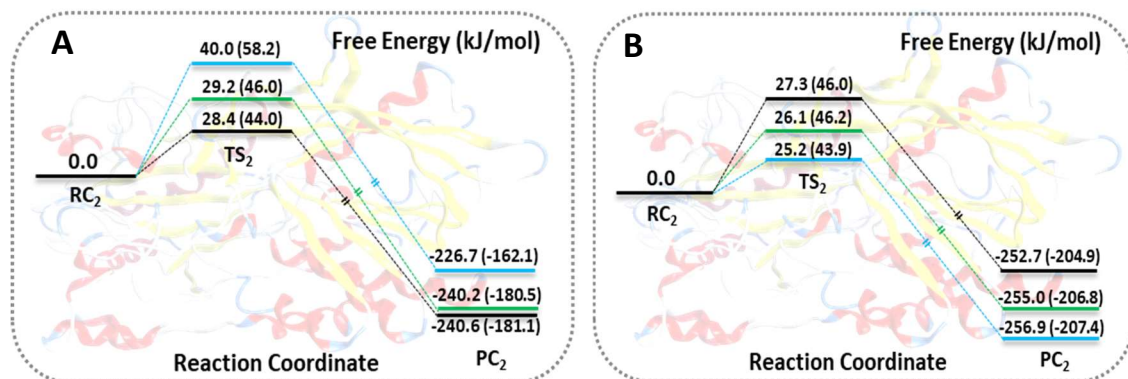
The obtained free energy surfaces for the elimination mechanism using two different functionals including M06 and M06L are shown in **Figure 7.6**. Additionally, the influence of increasing the basis set from 6-31G(d)  $\rightarrow$  6-311G(d,p) on the reaction pathway and optimized molecular structures have been studied. Except when applying the M06/6-311G(d,p) level of theory, improving the basis set has not noticeably impacted the calculated free energy of activation and free energy of the reaction.

Comparing the free activation energy calculated using these two functionals indicates a good agreement with an error of  $< \sim 3$  kJ/mol. Although it is found that the M06L functional decreases the free energy of the reaction by  $\sim 15$  kJ/mol with respect to the M06 functional. It is while geometrical parameters of the optimized **RC**<sub>2</sub>, **TS**<sub>2</sub>, and **PC**<sub>2</sub> from these functionals do not differ greatly.

It is also worth mentioning that there is a good agreement between the free energy of activation obtained using the QM-cluster and ONIOM(QM/MM) approaches (compare **Figure 7.6** and **7.3**).

Considering the electrostatic effects of the protein environment on the inner layer (QM region), the electrostatic embedding (EE) has been employed on the single-point energy calculation. As outlined in **Figure 7.6**, the electrostatic environment, regardless of the size of the basis set and applied functional, increases the free energy of activation and free energy of the reaction by  $\leq 20.1$  and  $\leq 64.6$  kJ/mol, respectively relative to the

corresponding values obtained using the mechanical embedding. This increase may be due to previously noted challenges with QM-region over-polarization.<sup>45</sup>



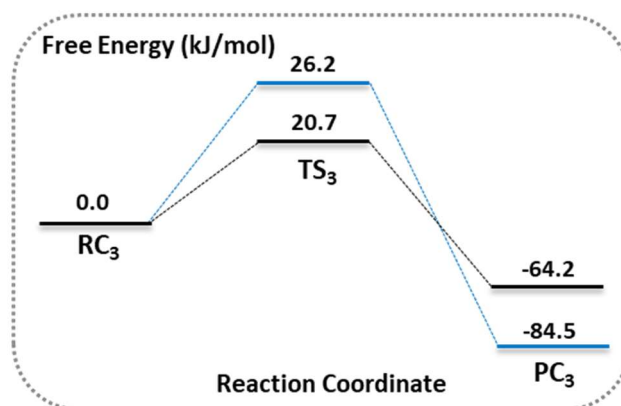
**Figure 7.6.** Calculated free energy surfaces using the ONIOM(QM/MM) approach; (A) M06 (B) M06L functionals, respectively. The green, black, and blue colored surfaces are for 6-31G(d), 6-31G(d,p), and 6-311G(d,p) level of optimization, respectively. Values in parentheses are calculated applying the electrostatic formalism on the single-point energy.

### 7.3.2 Mn-DddK Enzyme

Using the QM-cluster approach, the impact of Mn(II) metal ion on the DMSP decomposition activity was also studied. It is worth stating that optimized structures were obtained for the Mn(II)-containing bound active sites in their doublet, quartet, and sextet states. The sextet state was found to be lowest in energy. Hence, only results obtained for the sextet state are discussed herein.

By comparing the obtained optimized Ni(II) and Mn(II)-containing complexes along the reaction pathway, the main geometrical difference is observed in the Mn(II)-centered coordination in which by entering the DMSP into the enzyme active site, the octahedral metal center geometry of the unbound form is restructured into the partial octahedral metal center by sharing one coordination between the second carboxylate oxygen of Glu62 and one oxygen atom of substrate.

In contrast with the observed geometrical changes, the obtained free energy of activation of the studied elimination reaction represents a negligible change ( $< 1$  kJ/mol) comparing corresponding values using M06 and M06L functionals in the active site of Ni-DddK enzyme (compare **Figure 7.7** and **7.3**), although the changes in the free energy of reaction are slightly considerable ( $< 11$  kJ/mol).



**Figure 7.7.** Calculated free energy surfaces using the QM-cluster approach in Mn-DddK enzyme. The black and blue colored surfaces are for M06/6-31G(d,p) and M06L/6-31G(d,p) level of optimization, respectively.

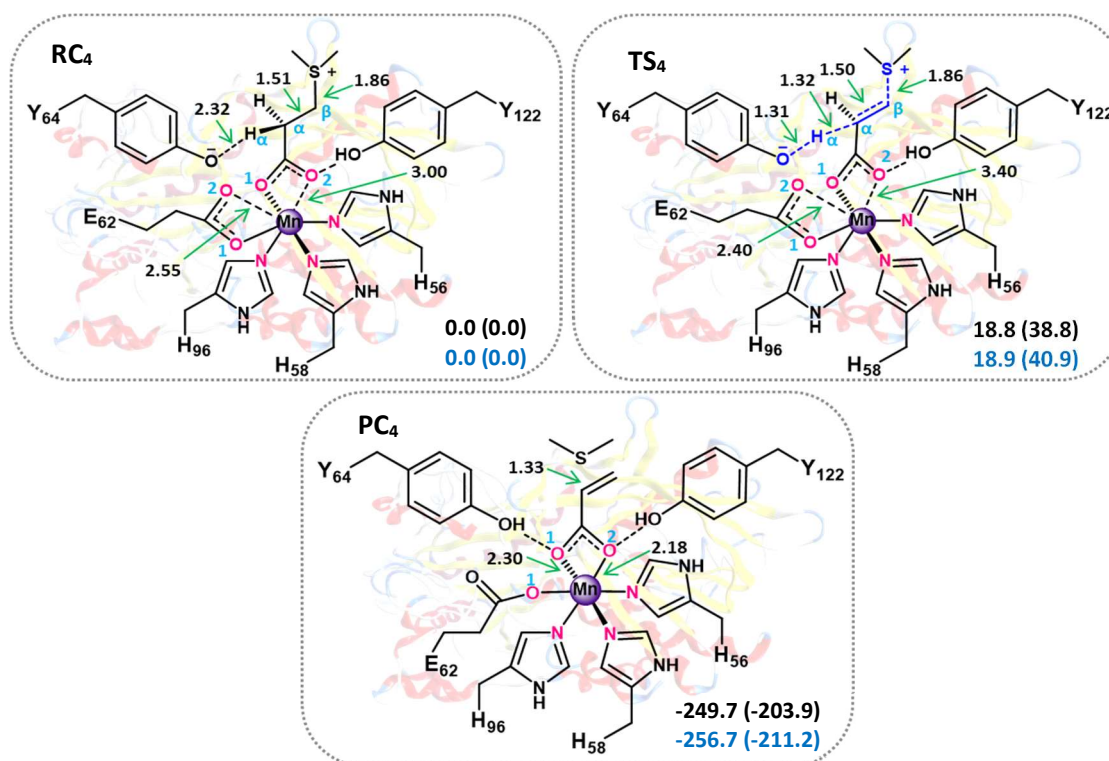
ONIOM(QM/MM) optimized complexes of reactive complex (**RC<sub>4</sub>**), transition structure (**TS<sub>4</sub>**), and product complex (**PC<sub>4</sub>**) at ONIOM(M06/6-31G(d,p);AMBER96) level of optimization are shown in **Figure 7.8**. As presented, in the bound active site of the enzyme, Mn(II) is ligated to His56, 58, and 96, and Glu62 in a bidentate fashion *via* two carboxylate oxygen atoms while DMSP is coordinated to the metal center (Mn(II)) in a monodentate fashion: ( $r_{\text{SubO}^1 \cdots \text{Mn}}$ : 2.03 Å *vs.*  $r_{\text{SubO}^2 \cdots \text{Mn}}$  is 3.00 Å). Although passing the transition structure at a cost of 18.8 kJ/mol and releasing DMS, acrylate is stabilized through bidentate coordination to Mn(II) and Glu62 rearranges to monodentate ligation to the metal center (see **Figure 7.8**).

In comparison with Ni(II)-centered coordination, the free energy of activation regardless of functionals applied (M06 and M06L), has been lowered by  $\leq 9.6$  kJ/mol. It is



while considering the electrostatic effect of the MM layer employing the electrostatic embedding, the observed changes in the free energy of activation have been decreased to  $\sim 5$  kJ/mol (compare **Figure 7.8** and 7.6).

Furthermore, the product complex (**PC<sub>4</sub>**) in the active site of Mn-centered coordination has been more stabilized with respect to the corresponding one in the active site of Ni-centered coordination.



**Figure 7.8.** QM/MM optimized structures of **RC<sub>4</sub>**, **TS<sub>4</sub>**, and **PC<sub>4</sub>** with selected bond distances in ångströms (Å) and the relative free energy of **TS<sub>4</sub>** and **PC<sub>4</sub>** with respect to the **RC<sub>4</sub>** in Mn-centered coordination. Black and blue colored values are related to M06 and M06L functionals, respectively. Values in parentheses are calculated utilizing the electrostatic formalism on the single point energy.

Comparing the obtained results of the decomposition mechanism indicates that the catalytic rate of DMSP cleavage in the active site of DddK enzyme in the presence of studied transition metal ions decreases in the following order: Mn(II) > Ni(II). The



observed order follows the normal rate of the water-exchange reaction on divalent transition metal ions in the first hydration layer:  $\text{Cu(II)} > \text{Cr(II)} > \text{Zn(II)} > \text{Mn(II)} > \text{Fe(II)} > \text{Co(II)} > \text{Ni(II)} > \text{V(II)}$ .<sup>46-47</sup>

#### 7.4 Conclusion

In the present study, using computational modeling approaches including QM-cluster and (ONIOM)QM/MM, we have investigated the bound and unbound active site and mechanism of DMSP lyase in Ni-Dddk enzyme from ubiquitous SAR11 marine bacteria.

The QM-cluster studies show that Ni(II) is coordinated in an octahedral fashion by three histidyls, a glutamyl, tyrosyl residues, and a water molecule which by the introduction of the substrate into the enzyme active site, the Ni-bound water and tyrosyl are displaced and DMSP binds in a bidentate fashion. Although our findings considering the protein environment using the ONIOM(QM/MM) approach indicate that upon addition of DMSP to the active site, along with the removal of metal-bound water and tyrosyl, DMSP coordinates in a monodentate fashion by one oxygen atom of its end carboxylate group. The detachment of the metal-coordinated tyrosyl and its removal from the first coordination sphere, makes it directly act as a Lewis base to initiate a concerted pathway of  $\beta$ -elimination reaction to release DMS. This “Tyr-switching” mechanism should be considered as the initial key process in the DMSP decomposition.

In spite of observed differences in the metal ion ligations of the binding active site in our QM-cluster and (ONIOM)QM/MM methods, a good agreement between the free energy of activation using these two methods was observed.

Furthermore, we have investigated the impact of the presence of Mn(II) on the catalytic rate of DMSP cleaving in the DddK enzyme. Comparing the free energy surfaces obtained using QM-cluster approach for Ni(II) and Mn(II)-centered coordination reveals an excellent agreement between the calculated free energy of activation and free energy of reaction, despite observed partial octahedral coordination for Mn(II) centered ion. While

employing the ONIOM(QM/MM) method and considering the environmental effects on the studied Mn-centered coordination and DMSP lyase, the free energy of activation decreases by  $\sim 10$  kJ/mol relative to Ni(II)-centered coordination. This observed order follows the normal rate of the water-exchange reaction on divalent transition metal ions in the first hydration sphere:  $\text{Mn(II)} > \text{Ni(II)}$ . However, our findings of the catalytic rate of DMSP decomposition does not follow the observed experimental order in the catalytic rate in the presence of different divalent transition metal ions in the active site of this enzyme. It shows the catalytic rate of DMSP lyase might be potentially influenced not only by the electronic effects of transition metal ions but also by some other features such as many long-range interactions and steric effects of protein environment, as some were observed using the QM/MM calculation, that need more detailed and deep investigations.

However, all of our findings and observation imply the importance of transition divalent metal ions acting as the cofactor in the DddK enzyme to facilitate the DMSP cleavage and DMS formation. Significantly, the results have broadened our understanding of function and structure of the newly-found DddK enzyme in the bacterial DMSP cleavage and DMS formation.

To assess the ability of *meta*-GGA functionals to describe the studied elimination mechanism and examining this metalloenzyme active site, two different functionals including M06 and M06L have been also employed in the current study. The results, despite the observed close agreement on describing the activation barrier and structural parameters, imply a better performance of M06L functional with no HF exchange contribution to describe transition metal ion-containing enzymes.

## 7.5 References

1. Raina, J.-B.; Tapiolas, D. M.; Forêt, S.; Lutz, A.; Abrego, D.; Ceh, J.; Seneca, F. O.; Clode, P. L.; Bourne, D. G.; Willis, B. L.; Motti, C. A., DMSP Biosynthesis by an Animal and its Role in Coral Thermal Stress Response. *Nature* **2013**, *502*, 677.

2. Otte, M. L.; Wilson, G.; Morris, J. T.; Moran, B. M., Dimethylsulphoniopropionate (DMSP) and related compounds in higher plants. *J. Exp. Bot.* **2004**, *55*, 1919-1925.
3. Kiene, R. P.; Linn, L. J.; Bruton, J. A., New and Important Roles for DMSP in Marine Microbial Communities. *J. Sea. Res.* **2000**, *43*, 209-224.
4. Stefels, J., Physiological Aspects of the Production and Conversion of DMSP in Marine Algae And Higher Plants. *J. Sea. Res.* **2000**, *43*, 183-197.
5. Sunda, W.; Kieber, D. J.; Kiene, R. P.; Huntsman, S., An antioxidant function for DMSP and DMS in marine algae. *Nature* **2002**, *418*, 317.
6. Strom, S.; Wolfe, G.; Holmes, J.; Stecher, H.; Shimeneck, C.; Lambert, S.; Moreno, E., Chemical defense in the microplankton I: Feeding and Growth Rates of Heterotrophic Protists on the DMS-Producing Phytoplankter *Emiliana Huxleyi*. *Limnol. Oceanogr.* **2003**, *48*, 217-229.
7. Wolfe, G. V.; Steinke, M.; Kirst, G. O., Grazing-Activated Chemical Defence in a Unicellular Marine Alga. *Nature* **1997**, *387*, 894-897.
8. Cao, H. Y.; Wang, P.; Xu, F.; Li, P. Y.; Xie, B. B.; Qin, Q. L.; Zhang, Y. Z.; Li, C. Y.; Chen, X. L., Molecular Insight into the Acryloyl-CoA Hydration by AcuH for Acrylate Detoxification in Dimethylsulfoniopropionate-Catabolizing Bacteria. *Front. Microbiol.* **2017**, *8*.
9. Van Alstyne, K. L.; Wolfe, G. V.; Freidenburg, T. L.; Neill, A.; Hicken, C., Activated defense systems in marine macroalgae: Evidence for an Ecological Role for DMSP Cleavage. *Mar. Ecol. Prog. Ser.* **2001**, *213*, 53-65.
10. Brummett, A. E.; Schnicker, N. J.; Crider, A.; Todd, J. D.; Dey, M., Biochemical, Kinetic, and Spectroscopic Characterization of *Ruegeria Pomeroyi* DddW-A Mononuclear Iron-Dependent DMSP Lyase. *Plos One* **2015**, *10*.
11. Bullock, H. A.; Luo, H. W.; Whitman, W. B., Evolution of Dimethylsulfoniopropionate Metabolism in Marine Phytoplankton and Bacteria. *Front. Microbiol.* **2017**, *8*, 637.

12. Moran, M. A.; Reisch, C. R.; Kiene, R. P.; Whitman, W. B., Genomic Insights into Bacterial DMSP Transformations. *Annu. Rev. Mar. Sci.* **2012**, *4*, 523-542.
13. Giovannoni, S. J., SAR11 Bacteria: The Most Abundant Plankton in the Oceans. *Annu. Rev. Mar. Sci.* **2017**, *9*, 231-255.
14. Lovelock, J. E.; Maggs, R. J.; Rasmussen, R. A., Atmospheric Dimethyl Sulphide and the Natural Sulphur Cycle. *Nature* **1972**, *237*, 452.
15. Andreae, M. O., Ocean-Atmosphere Interactions in the Global Biogeochemical Sulfur Cycle. *Mar. Chem.* **1990**, *30*, 1-29.
16. Seymour, J. R.; Simo, R.; Ahmed, T.; Stocker, R., Chemoattraction to Dimethylsulfoniopropionate Throughout the Marine Microbial Food Web. *Science* **2010**, *329*, 342-345.
17. Raina, J. B.; Dinsdale, E. A.; Willis, B. L.; Bourne, D. G., Do the Organic Sulfur Compounds DMSP and DMS Drive Coral Microbial Associations? *Trends Microbiol.* **2010**, *18*, 101-108.
18. Raina, J. B.; Tapiolas, D.; Willis, B. L.; Bourne, D. G., Coral-Associated Bacteria and Their Role in the Biogeochemical Cycling of Sulfur. *Appl. Environ. Microb.* **2009**, *75*, 3492-3501.
19. Malin, G., Oceans - New Pieces for the Marine Sulfur Cycle Jigsaw. *Science* **2006**, *314*, 607-608.
20. Simo, R., Production of Atmospheric Sulfur by Oceanic Plankton: Biogeochemical, Ecological and Evolutionary Links. *Tends Ecol. Evol.* **2001**, *16*, 287-294.
21. Reisch, C. R.; Moran, M. A.; Whitman, W. B., Bacterial catabolism of Dimethylsulfoniopropionate (DMSP). *Front. Microbiol.* **2011**, *2*, 172.
22. Roy, K., Thiols and Organic Sulfides. In *Ullmann's Encyclopedia of Industrial Chemistry*, 2000.
23. Li, C. Y.; Wei, T. D.; Zhang, S. H.; Chen, X. L.; Gao, X.; Wang, P.; Xie, B. B.; Su, H. N.; Qin, Q. L.; Zhang, X. Y.; Yu, J.; Zhang, H. H.; Zhou, B. C.; Yang, G. P.; Zhang, Y.

Z., Molecular Insight Into Bacterial Cleavage of Oceanic Dimethylsulfoniopropionate Into Dimethyl Sulfide. *Proc. Natl. Acad. Sci. U.S.A* **2014**, *111*, 1026-1031.

24. Moran, M. A.; Reisch, C. R.; Kiene, R. P.; Whitman, W. B., Genomic Insights into Bacterial DMSP Transformations. In *Annual Review of Marine Science, Vol 4*, Carlson, C. A.; Giovannoni, S. J., Eds. 2012; Vol. 4, pp 523-542.

25. Brummett, A. E.; Dey, M., New Mechanistic Insight from Substrate- and Product-Bound Structures of the Metal-Dependent Dimethylsulfoniopropionate Lyase DddQ. *Biochemistry* **2016**, *55*, 6162-6174.

26. Curson, A. R. J.; Sullivan, M. J.; Todd, J. D.; Johnston, A. W. B., DddY, a Periplasmic Dimethylsulfoniopropionate Lyase Found in Taxonomically Diverse Species Of Proteobacteria. *Isme J.* **2011**, *5*, 1191-1200.

27. Todd, J. D.; Curson, A. R. J.; Kirkwood, M.; Sullivan, M. J.; Green, R. T.; Johnston, A. W. B., DddQ, A Novel, Cupin-Containing, Dimethylsulfoniopropionate Lyase in Marine Roseobacters and in Uncultured Marine Bacteria. *Environ. Microbiol.* **2011**, *13*, 427-438.

28. Alcolombri, U.; Ben-Dor, S.; Feldmesser, E.; Levin, Y.; Tawfik, D. S.; Vardi, A., Identification of the Algal Dimethyl Sulfide-Releasing Enzyme: A Missing Link in the Marine Sulfur Cycle. *Science* **2015**, *348*, 1466-1469.

29. Lei, L.; Cherukuri, K. P.; Alcolombri, U.; Meltzer, D.; Tawfik, D. S., The Dimethylsulfoniopropionate (DMSP) Lyase and Lyase-Like Cupin Family Consists of Bona Fide DMSP Lyases as Well as Other Enzymes with Unknown Function. *Biochemistry* **2018**, *57*, 3364-3377.

30. Wang, P.; Chen, X. L.; Li, C. Y.; Gao, X.; Zhu, D. Y.; Xie, B. B.; Qin, Q. L.; Zhang, X. Y.; Su, H. N.; Zhou, B. C.; Xun, L. Y.; Zhang, Y. Z., Structural and Molecular Basis for the Novel Catalytic Mechanism and evolution of DddP, an abundant peptidase-like bacterial Dimethylsulfoniopropionate lyase: a new enzyme from an old fold. *Mol. Microbiol.* **2015**, *98*, 289-301.

31. Wang, Y.; Chen, S. L., How is DMSP Decomposed When Catalyzed by RlDddP Binuclear Iron DMSP Lyase? *J. Catal.* **2018**, *360*, 1-8.
32. Schnicker, N. J.; De Silva, S. M.; Todd, J. D.; Dey, M., Structural and Biochemical Insights into Dimethylsulfoniopropionate Cleavage by Cofactor-Bound DddK from the Prolific Marine Bacterium *Pelagibacter*. *Biochemistry* **2017**, *56*, 2873-2885.
33. Vraspir, J. M.; Butler, A., Chemistry of Marine Ligands and Siderophores. *Annu. Rev. Mar. Sci.* **2009**, *1*, 43-63.
34. Pribula, A., The Elements, Their Origin, Abundance, and Distribution (Cox, P.A.). *J. Chem. Educ.* **1991**, *68*, A112.
35. *Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group ULC.*, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2018.
36. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K., Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781-1802.
37. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; N. Rega; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; J. V. Ortiz; Cioslowski, J.; Fox, D. J. *Gaussian 09*, E.01; Gaussian, Inc.: Wallingford CT, 2015.

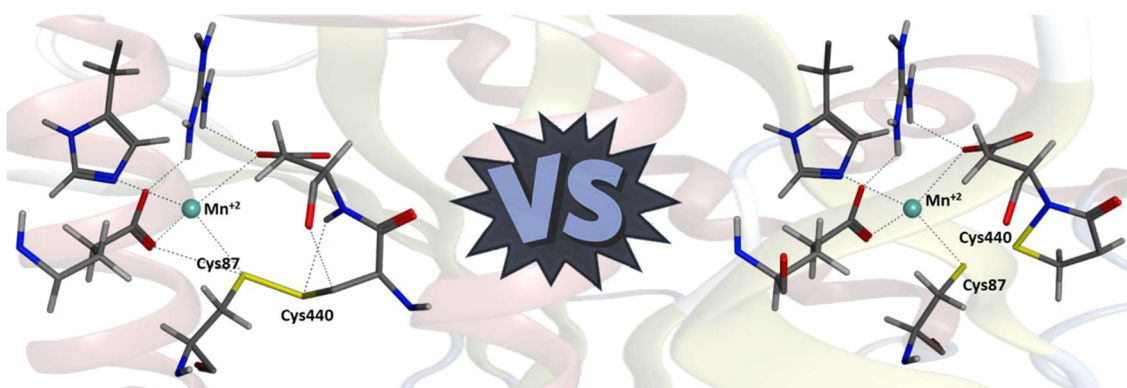
38. Himo, F., Recent Trends in Quantum Chemical Modeling of Enzymatic Reactions. *J. Am. Chem. Soc.* **2017**, *139*, 6780-6786.
39. Zhao, Y.; Truhlar, D. G., The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215-241.
40. Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H. B.; Frisch, M. J., Geometry Optimization with QM/MM, ONIOM, and Other Combined Methods. I. Microiterations and Constraints. *J. Comput. Chem.* **2003**, *24*, 760-769.
41. Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z. F.; Liu, F. Y.; Li, H. B.; Ding, L. N.; Morokuma, K., The ONIOM Method and Its Applications. *Chem. Rev.* **2015**, *115*, 5678-5796.
42. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1996**, *118*, 2309-2309.
43. Park, H. I.; Ming, L.-J., The Mechanistic Role of the Coordinated Tyrosine in Astacin. *J. Inorg. Biochem.* **1998**, *72*, 57-62.
44. Bas, D. C.; Rogers, D. M.; Jensen, J. H., Very Fast Prediction and Rationalization of pKa Values for Protein–Ligand Complexes. *Proteins: Struct. Funct. Bioinf.* **2008**, *73*, 765-783.
45. Groenhof, G., Introduction to QM/MM Simulations. In *Biomolecular Simulations: Methods and Protocols*, Monticelli, L.; Salonen, E., Eds. Humana Press: Totowa, NJ, 2013; pp 43-66.
46. Helm, L.; Merbach, A. E., Water Exchange on Metal Ions: Experiments and Simulations. *Coordin. Chem. Rev.* **1999**, *187*, 151-181.

47. Akesson, R.; Pettersson, L. G. M.; Sandstroem, M.; Siegbahn, P. E. M.; Wahlgren, U., Theoretical Study of Water-Exchange Reactions for the Divalent Ions of the First Transition Period. *J. Phys. Chem.* **1993**, *97*, 3765-3774.



## CHAPTER 8

# Computational Insights into A Potential Disulfide-Sulfenylamide Exchange Mechanism in 3-Deoxy-d-Arabino-Heptulosonate 7-Phosphate Synthase (DAH7PS)



## 8.1 Introduction

Within cells and organisms many proteins expand or alter their chemical functionality, provided *via* their constituent proteinogenic amino acids, through post-translational modifications (PTMs). A vast and diverse array of such modifications have been observed *in vivo*, *e.g.*, oxidation, SUMOylation, phosphorylation.<sup>1-4</sup> In addition, they may be reversible or irreversible, and are now known to often have physiologically essential roles such as cell signaling, and protein localization and regulation.<sup>5-7</sup> For example, in SUMOylation a small ubiquitin-like modifier (SUMO) peptide is attached to a protein. This modification has been identified as being involved in a variety of physiologically significant processes including, for example, apoptosis (cell death).<sup>8-9</sup>

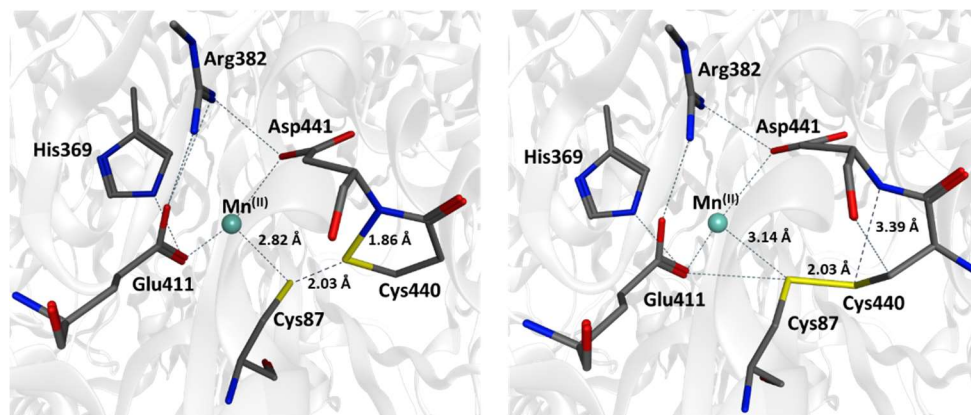
Whilst in principle any constituent amino acid residue in a protein may be modified, certain residues are more amenable to modification. In particular, cysteine residues are a key target for PTMs with cystine (disulfide) bonds being the most common covalent cross-link within proteins.<sup>10-11</sup> In addition, however, they are also known to undergo a range of covalent and often reversible modifications including nitrosylation, glutathionylation, persulfidation, and oxidation and that such resulting PTMs play physiologically important roles.<sup>12-17</sup> Amongst the plethora of known and possible cysteinyl modifications, there is a quickly increasing interest in those containing S—N bonds. This has been spurred on in part due to a number of recent discoveries including, for example, sulfilimine cross-links in collagen fibers and a recognition of their importance in tissue formation and function.<sup>18</sup> However, the mode of formation and the role of some experimentally observed cysteinyl modifications remains unknown or unclear.

For example, it has been observed that in several proteins (so far to date) the sulfur of a cysteinyl residue is able to react with the backbone amide bond nitrogen of its adjacent residue to form a cyclic sulfenylamide derivative.<sup>19-23</sup> Arguably the first recognized example of such a PTM with an identified enzymatically important role was in a Protein tyrosine phosphatase 1B (PTP1B).<sup>24</sup> It is a member of the large PTP enzyme family of

enzymes and which perform, for example, key roles in signaling pathways through tyrosine de/phosphorylation.<sup>25</sup> The subclass PTP1B in particular is important for insulin and leptin signaling.<sup>25-28</sup> In the case of PTP1B, the sulfenylamide is thought to be one of the few proven mechanisms, experimentally<sup>24</sup> and computationally<sup>29</sup>, by which the active site cysteinyl is protected from irreversible oxidative damage.<sup>22, 30-31</sup> More specifically, oxidation of the active site cysteinyl results in formation of a sulfenic acid (RSOH) derivative. Meanwhile, a histidyl and glutamyl residue form hydrogen bonds with the backbone amide carbonyl oxygen and -NH- groups, respectively. As a result, the amide bond is able to interconvert between its amide and iminol forms. In the latter form, the sulfenic acid's sulfur is able to nucleophilically attack nitrogen and form the sulfenylamide. Other computational investigations suggested that the formation of such a species may also be facilitated by structural constraints involving the protein backbone.<sup>23</sup> Furthermore, in a survey of available crystal structures in the Protein DataBank (PDB) at that time, they noted that four proteins including PTP1B, appeared to contain a sulfenylamide moiety.

In particular, as they noted, such a PTM also appears in an X-ray crystal structure (PDB ID: 3NUE) of the *Mycobacterium tuberculosis* metalloenzyme 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase (DAH7PS) complexed with tryptophan.<sup>32</sup> This enzyme catalyzes the critical first step in the Shikimate pathway of aromatic amino acid biosynthesis; the stereospecific aldol condensation of phosphoenolpyruvate (PEP) with D-erythrose-4-phosphate (E4P) to give 3-deoxy-D-*arabino*-heptulosonate-7-phosphate (DAH7P).<sup>32-35</sup> More specifically, within the active site a Mn(II) ion is ligated in part *via* an active site cysteinyl (Cys87) sulfur, while a second cysteinyl (Cys440) lies spatially close. In the X-ray crystal structure, however, Cys87 and Cys440 appear to be involved in two distinctly different bonding positions (**Figure 8.1**). In one conformation, the sulfur of Cys440 appears to be covalently bonded to the backbone nitrogen center of the adjacent residue Asp441 with an  $\text{Cys440S} \cdots \text{N}_{\text{Asp441}}$  distance of 1.86 Å. Concomitantly, however, the sulfur of Cys87 ( $\text{Cys87S}$ ) is positioned 2.03 Å from the  $\text{Cys440S}$  and 2.82 Å from the active

site Mn(II) ion. In contrast, in the second conformation, the  $\text{Cys440S}\cdots\text{N}_{\text{Asp441}}$  distance is significantly greater at 3.39 Å; and thus, there is no longer a covalent bond between these two centers. Concomitantly, the  $\text{Cys87S}\cdots\text{Mn(II)}$  distance has also increased markedly to 3.14 Å, suggesting that they are more weakly interacting in this second conformation. However, the  $\text{Cys87S}\cdots\text{S}_{\text{Cys440}}$  distance remains at 2.03 Å. There are known challenges with X-ray crystal structures and their relationship to their *in vivo* structures such as crystal packing effects, the differing environments, *etc.* The above observations in X-ray crystal structure of DAH7PS, however, appear to suggest that two conformations are possible; (i)  $\text{Cys87S}$  is ligated to the Mn(II) ion and Cys440 has formed a cyclic sulfenylamide with Asp441 (the cyclic sulfenylamide conformation), and (II) where Cys87 and Cys440 have instead formed a disulfide bond with each other and no longer interact with the Mn(II) nor make a sulfenylamide (the disulfide conformation). What role such a possible conformational switch may also have on the enzyme's activity is unknown. To the best of our knowledge such an observed disulfide–sulfenylamide switching mechanism would be the first such recognized occurrence of this type of shuttling reaction.



**Figure 8.1.** Illustration of two different active sites of DAH7P enzyme (PDB ID: 3NUE) with cyclic sulfenylamide (on the left) and disulfide bond (on the right) and selected key distances in ångstroms (Å).

In this present study, we have used a multi-scale computational approach to examine the active site of 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase (DAH7PS) from *Mycobacterium tuberculosis*. More specifically, we have complementarily applied molecular dynamics (MD) simulations, QM-cluster and quantum mechanics/molecular mechanics (QM/MM) approaches to examine the potential interconversion between the cyclic sulfenylamide and disulfide conformations. As part of this study, we have considered the role of other residues and/or aqueous solution on such a process.

### 8.2 Computational Methods

#### 8.2.1 Molecular Dynamics (MD) Simulations

The Molecular operating environment (MOE) software was used to prepare all chemical models for MD simulations.<sup>36</sup> The X-ray crystal structure of 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase from *Mycobacterium tuberculosis* complexed with tryptophan (PDB ID: 3NUE)<sup>32</sup> with a resolution of 2.5 Å was as the initial template structure.

Two different chemical models were then prepared for simulations. In one a cyclic sulfenylamide involving Cys440 and Asp441 was formed, as observed in the X-ray crystal structure, while in the other a disulfide bond was formed involving the spatially adjacent sulfurs of Cys87 and Cys440. Protonation states of all residues were assigned based on the default PROPKA protonation tool available *via* MOE. The models were solvated to the depth of a 6 Å layer of water around the entire system, and then minimized using the AMBER14:EHT force field. It is noted that the applicability of the AMBER forcefield for Mn-containing metalloproteins was previously considered by Neves *et al.*<sup>37</sup> Each optimized structure was then equilibrated for 100 ps with the temperature being increased from 0 to 300 K, at constant pressure. This was followed for each complex by an unconstrained 5 ns MD simulations with a time step of 2 fs, under constant pressure and temperature, using the NAMD program (default settings).<sup>38</sup>

### 8.2.2 QM-Cluster and QM/MM Calculations

To further elucidate the potential disulfide-sulfenylamide exchange mechanism, the QM-cluster and hybrid ONIOM(QM/MM) approaches were used. For all such studies, including optimizations and harmonic vibrational frequencies calculations, the Gaussian09 software package was used.<sup>39</sup>

The QM-cluster approach has been successfully applied to the study of numerous enzymatic mechanisms and processes, especially those involving a metal ion.<sup>40-42</sup> In the present study, it has been used to gain additional insights into the possible exchange mechanism. The chemical model used for the QM-cluster studies consisted of the Mn(II) ion and at least the side chains of those residues to which it is coordinated, as well as those that might participate in the exchange mechanism: Asp441 (entire residue), His369, Glu411, Cys87, Cys440 (entire residue), and Tyr436. The rest of the protein environment also omitted after capping each residue at its C $\alpha$  carbon atoms by addition of hydrogen atoms. These carbon atoms were also held fixed in position to ensure the structural integrity of the chemical model was maintained. The density functional theory method B3LYP<sup>43-44</sup> was used as it has been shown to perform well for Mn-containing systems.<sup>45-48</sup> Additionally, to assess basis set performance, a systematic series of basis sets including 6-31G(d), 6-311G(d), 6-31G(d,p), and 6-311G(d,p) were used. Furthermore, to more accurately model the effects of non-covalent and dispersion interactions the empirical D3 dispersion correction by Grimme was also used in combination with the B3LYP/6-311G(d) and B3LYP/6-311G(d,p) levels of theory.<sup>49</sup>

The hybrid ONIOM(QM/MM) approach was also applied so as to accurately consider the influence of the protein environment on the possible exchange mechanism. The whole chemical system was divided into two regions (a QM and an MM region) based on their contributions to the mechanism. The QM region included the Mn(II) ion and at least the side chain groups of the key residues in the enzyme active site to which it is ligated (the complete Asp441 residue was included), as well as Cys440, and the side chain of Tyr436.

Based on results obtained for the above QM-cluster studies, this QM subsystem was described using B3LYP and B3LYP-D3 in conjunction with the 6-31G(d,p) basis set. Meanwhile, the remaining part of the protein was placed in the low layer (MM region) and treated molecular mechanically using the AMBER96 forcefield. Moreover, in the MM layer, the C $\alpha$  centers were held fixed in position while the rest were free to move during optimizations. All the ONIOM(QM/MM) calculations were carried out with the mechanical embedding (ME) formalism.

Several alternate ONIOM(QM/MM) models were also used to examine additional aspects that may influence the possible mechanism or the reliability of the results. More specifically, a second model was created in which the side chain of Thr438 was included in the QM region in order to examine its potential influence on the exchange mechanism. For this model, the QM region was described using the B3LYP/6-31G(d,p) method, while the rest was again described molecular mechanically using AMBER96. A second alternate model was also used to gain further insights into the effects of the surrounding protein environment and changes in coordination of the Mn(II) ion. Specifically, the QM-region was expanded to include the Mn(II), His369, Glu411, Asp441, Cys87 and 440, Pro442, Thr438, Ala439, and Tyr436 and a number of explicit water molecules (6 in total). This larger QM-layer was again described using the B3LYP/6-31G(d,p) level of theory, while the rest of the chemical complex (the MM-layer) was again described using the AMBER96 forcefield. Finally, the effect of improving the basis set from 6-31G(d,p) to 6-311G(d,p), for the Cys87 and Cys440 sulfur atoms in the QM-region only, was also examined. It is noted that all other atoms in the QM-region were described using the 6-31G(d,p) basis set as before, and the MM-region was also described as before.

### 8.3 Results and Discussion

#### 8.3.1 QM-Cluster Investigation

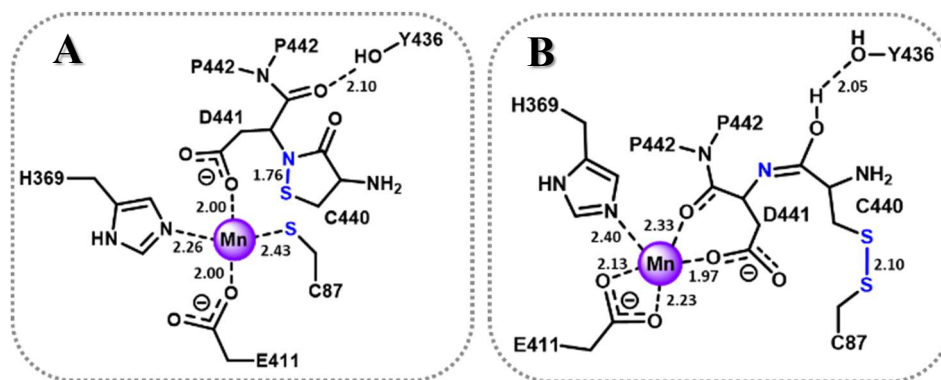
To study the exchange mechanism and gain the initial geometrical insight into the metal ion center and its ligation, the active site of enzyme involving the cyclic sulfenylamide result of thiolate of Cys440 and amide nitrogen of Asp441 has been optimized applying B3LYP functional with a wide range of basis sets (see methodology section). **Figure 8.2A** indicates the obtained optimized structure using B3LYP/6-31G(d,p). As shown in the active site of the enzyme, Mn(II) is ligated to His369, deprotonated Cys87, Glu411, and Asp441 in a monodentate fashion. Geometrical analysis of optimized structures using different basis sets has represented an insignificant change in the measured  $r(\text{C}_{440}\text{S}—\text{N}_{\text{D441}})$  in the sulfenylamide ring and  $r(\text{C}_{87}\text{S}\cdots\text{Mn})$ . The only exception was observed considering the dispersion interaction applying 6-311G(d)-D3 in which  $r(\text{S}\cdots\text{Mn})$  has been lengthened by 0.04 Å relative to other applied basis sets. Additionally, in the obtained optimized structures using 6-31G(d) and 6-311G(d)-D3,  $\text{Y}_{436}\text{OH}$  has shown a strong H-bond with the oxygen atom of  $\text{C}_{440}\text{C}=\text{O}$ ;  $r(\text{H}\cdots\text{O})$ : 1.78 Å. In contrast, applying other basis sets, this type of interaction has been replaced with a weaker H-bond between  $\text{D}_{441}\text{O}$  and  $\text{Y}_{436}\text{OH}$ ;  $r(\text{H}\cdots\text{O})$ : 2.10-2.20 Å (**Figure 8.2A**)

To study the disulfide-sulfenylamide exchange mechanism in the enzyme active site,  $\text{C}_{87}\text{S}—\text{S}_{\text{C440}}$  bond formation was explored through protonating the oxygen atom of cyclic sulfenylamide ( $\Delta E = -1368.8$  kJ/mol). As **Figure 8.2B** represents in the obtained optimized structure Cys87 has no longer coordination with Mn(II). It forms a disulfide bond with a bond length of 2.10 Å with Cys440 ( $\text{C}_{87}\text{S}—\text{S}_{\text{C440}}$ ), coincident with the sulfur-nitrogen bond ( $\text{C}_{440}\text{S}—\text{N}_{\text{D441}}$ ) cleavage.

A result of this switching mechanism, metal ion coordination has been replaced with two new ligations; second oxygen atom of the carboxylate group of Glu411 and oxygen atom of the amide bond in Asp411. Hence in the obtained active site, His369, Glu411 in a bidentate fashion through oxygen atoms of its carboxylate group and Asp411 through one



of its oxygen atom of carboxylate and oxygen atom of its amide group are ligated to the Mn(II) to form a penta-coordinated metal ion complex.



**Figure 8.2.** Schematic representation of optimized structures of the active site (A) and protonated oxygen amide in the active site (B) with the selected bond distances shown in ångstroms (Å) at the B3LYP/6-31G(d,p) level of theory.

Considering the influence of Tyr436 in the second sphere of Mn(II) coordination on the switching mechanism of disulfide-sulfenylamide and acting as a proton donor, proton transferring process from the hydroxyl group of Tyr436 to the oxygen atom of cyclic sulfenylamide alongside the S—S bond formation was also studied. All the obtained optimized structures from this proton transferring and switching mechanism imply the cyclic sulfenylamide and neutral Tyr436 preference over disulfide bond formation since the Cys87 favors metal ion coordination and proton transferred back to Tyr436.

With the initially obtained insight into the switching mechanism and with the aim of taking the protein environment into account and gaining a deeper understanding of disulfide-sulfenylamide exchange mechanism and other influential roles of the local biochemical environment, we have expanded the models into the ONIOM(QM/MM) calculations.

### 8.3.2 QM/MM Investigation; Small Model of QM Region

The QM/MM optimized structure of the active site at ONIOM(B3LYP/6-31G(d,p):AMBER96) level of optimization, with selected distances in ångstroms (Å) is shown in **Figure 8.3C**. As it is obvious, in this optimized structure Mn(II) is ligated to Glu411 in a bidentate fashion through both oxygen atoms of its carboxylate group, His396, Asp441 in a monodentate fashion and Cys87 leading to a penta-coordinated metal center. All coordinated ligands are 1.99-2.59 Å from the Mn(II) center. Thiolate of Cys440 is also bonded to the amide nitrogen of Asp441 to form a cyclic sulfenylamide. Obviously considering the protein environment has influenced the metal-center coordination remarkably compared to the obtained structure from the QM-cluster approach; 5-ligated Mn(II) vs. 4-ligated Mn(II) results from bidentate-coordinated Glu411 in the QM/MM calculation vs. monodentate-coordinated Glu411 in the QM-cluster approach. Additionally, the protein environment has lengthened the  $r_{(C_{87}S \cdots Mn)}$  by 0.16 to 2.59 Å. Although the change in  $C_{440}S-N_{D441}$  is less significant; increasing by only 0.05 Å. As well the inclusion of dispersion interaction effects *via* the use of Grimme's correction (B3LYP-D3) has negligible effects on the obtained geometrical features, except an H-bonding between  $C_{440}O-H_{Y436}$  which was not observed using B3LYP.

Taking the protein environment into account, the influence of protonated oxygen atom of the sulfenylamide bond on the exchange mechanism has been also studied ( $\Delta E = -1534.3$  kJ/mol). As shown in **Figure 8.3D**, this protonating process causes Cys87 thiolate to leave the metal ion first coordination sphere and form a disulfide bond with Cys440 ( $C_{87}S-S_{C440}$ ) which is coincident with  $C_{440}S-N_{D441}$  bond cleavage. In contrast with obtained results from corresponding QM-cluster approach, removing the ligated Cys87 from metal ion coordination has not been replaced with any other coordination instead the remained ligands make stronger interactions with the metal ion.

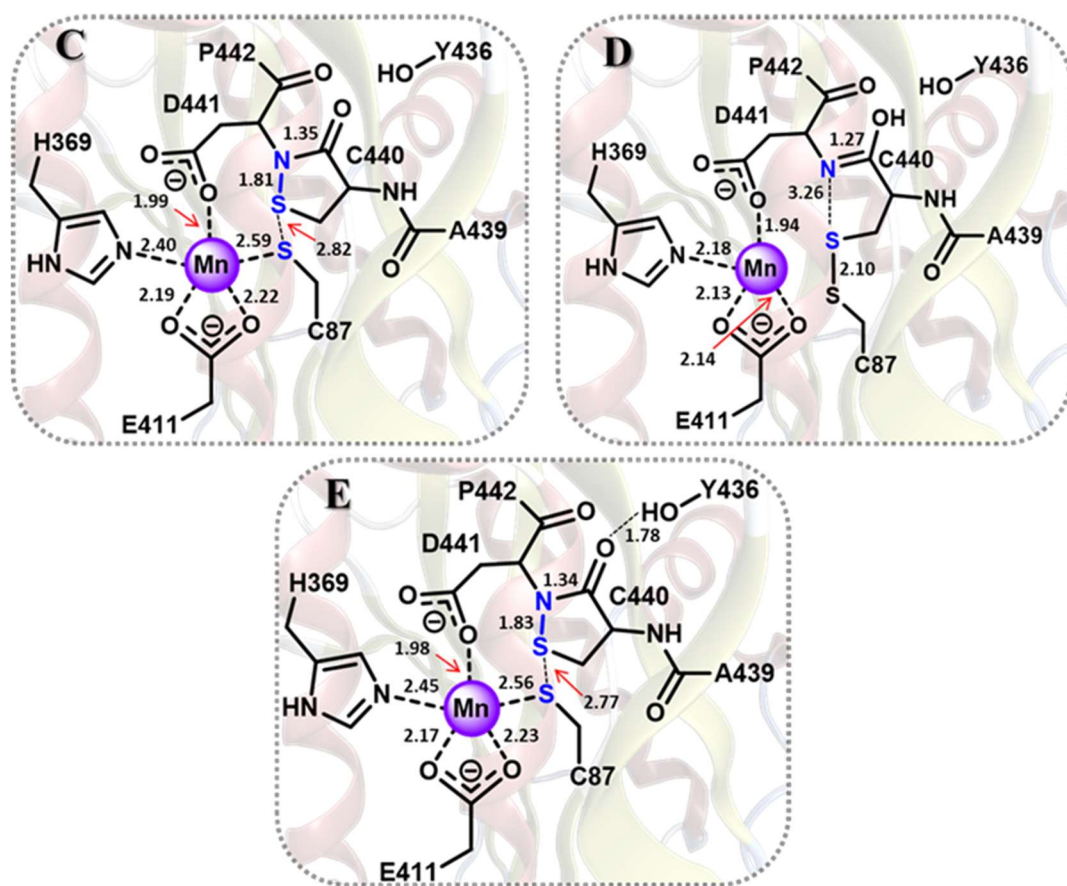
To assess how the Tyr436 might behave differently acting as a proton donor from QM-cluster approach, its influence on the exchange mechanism also studied applying the

ONIOM(QM/MM) method. To this aim, proton transferring process from the hydroxyl group of Tyr436 to the oxygen atom of the sulfenylamide bond, simultaneous with removing Cys87 thiolate from metal coordination and  $C_{87}S-S_{C440}$  bond formation was explored. Like the QM-cluster approach, obtained optimized complex reflects the S—N preference over  $C_{87}S-S_{C440}$ , which is coincident with the proton transferring back to deprotonated Tyr436 and Cys87 thiolate ligation to Mn(II). Although the obtained optimized complex is 314.6 kJ/mol lower in energy than the corresponding optimized structure from the initial crystal structure, no significant geometrical change in Mn(II) coordination was observed, except the strong H-bond between  $C_{440}O \cdots H_{Y436}$  with a distance of 1.78 Å (compare **Figure 8.3C** and **8.3E**).

Dispersion interaction effects on this process *via* Grimme's correction (B3LYP-D3) were also investigated. The results do not show the considerable changes in the obtained geometrical parameters relative to the corresponding optimized structures using B3LYP.

### 8.3.2.1 Mn(II) Influence on the Exchange Mechanism

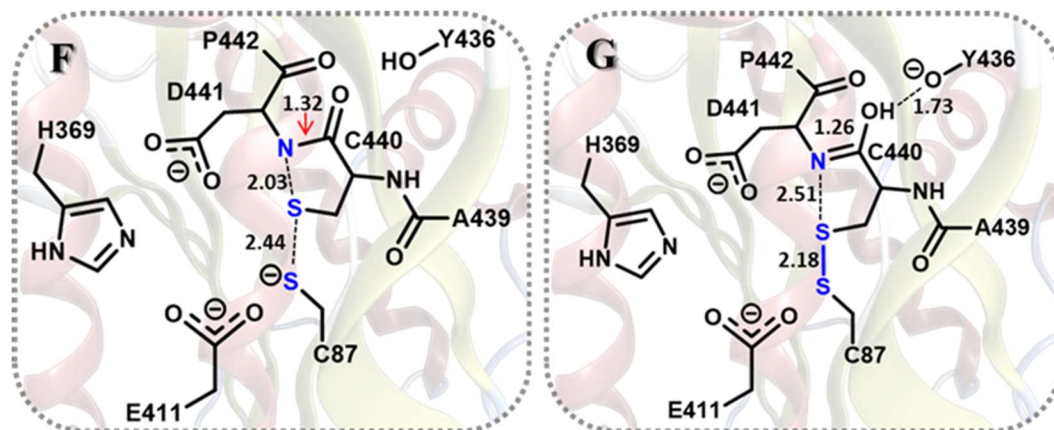
To investigate the Mn(II) impact on the exchange mechanism, the enzyme active site has also been optimized employing ONIOM(B3LYP/6-31G(d,p):AMBER96) in the absence of metal ion. As it is shown in **Figure 8.4F**, manganese absence in the active site leads to  $C_{440}S-N_{D441}$  bond distance increasing by 0.22 from 1.81 to 2.03 Å. On the other hand,  $r(C_{87}S \cdots S_{C440})$  distance indicates a decrease of 0.38 Å from 2.82 to 2.44 Å (compare **Figure 8.3C** and **8.4F**). It is worth mentioning that considering dispersion interaction effects are more obvious in the active site without a metal ion compared to the active site with the metal ion. In fact, dispersion interaction causes  $C_{440}S-N_{D441}$  to be lengthened to 2.38 Å. Additionally, the formation of a covalently stronger interaction of  $C_{87}S-S_{C440}$  with a bond length of 2.20 Å is observed using the Grimme's correction (B3LYP-D3).



**Figure 8.3.** Schematic illustration of optimized structures of the active site (C) and protonated oxygen amide in the active site (D), and a complex result of examining proton transferring from Tyr436 to oxygen atom of sulfenylamide bond (E) with the selected distances in ångstroms (Å) using ONIOM(B3LYP/6-31G(d,p):AMBER96).

Furthermore, the influence of Mn(II) absence in the active site on the proton transferring process from the hydroxyl group of Tyr436 to the oxygen atom of cyclic sulfenylamide and switching mechanism have also been explored. Comparison with corresponding QM-cluster and QM/MM models with the metal ion in the active site and proton transferring process, in which  $C_{440}S-N_{D441}$  and the neutral Tyr436 are favored over  $C_{440}S-S_{C87}$  and the deprotonated Tyr436, proton transferring while Mn(II) is absent in the active site results in disulfide bond formation ( $C_{87}S-S_{C440}$ ) with a distance of a 2.18 Å (see **Figure 8.4G**). Also, this proton transferring leads to increasing  $r(C_{440}S \cdots N_{D441})$  from 2.03 to 2.51 Å. This

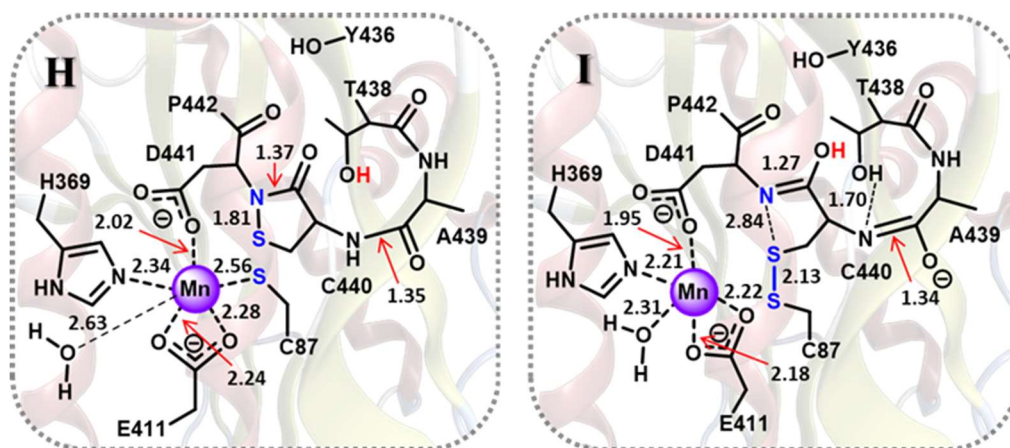
obtained optimized structure lies 142.1 kJ/mol lower in energy than the former structure without Mn(II). The inclusion of dispersion interaction effects *via* the Grimme's correction represents a negligible impact on the  $r_{(C440S\cdots N_{D411})}$  and  $r_{(C87S-S_{C440})}$  in the obtained optimized complex ( $< 0.02$  Å).



**Figure 8.4.** Schematic representation of optimized complexes of the active site (F) and the result of proton transferred from Tyr to oxygen amide in the active site (G) without metal ion with selected distances shown in ångstroms (Å) using ONIOM(B3LYP/6-31G(d,p):AMBER96).

### 8.3.2.2 Proton Transferring from Threonine

To assess the Thr438 contribution, positioned at neighboring cyclic sulfenylamide to the exchange mechanism, proton transferring process from the hydroxyl group of the threonine side chain to the oxygen atom of cyclic sulfenylamide has also been investigated. **Figure 8.5** represents two obtained optimized complexes of the crystal structure including Thr438 (**Figure 8.5H**) and the complex resulting from proton transferring process (**Figure 8.5I**). As shown, this proton shuttling leads to  $C440S-N_{D441}$  bond cleavage and  $C87S-S_{C440}$  bond formation ( $\Delta E = -672.3$  kJ/mol). Interestingly, this proton transferring is coincident with the second proton shuttling from Cys440 nitrogen amide to the side chain oxyanion of Thr438;  $_{Thr}PA = 1338.4$  kJ/mol.



**Figure 8.5.** Schematic representation of optimized complexes of the active sites including Thr438 (**H**) and the complex resulted from proton transferring from Thr438 to oxygen amide (**I**) with the selected distances in ångstroms (Å) using ONIOM(B3LYP/6-31G(d,p):AMBER96).

As a result of the above observations, presence of hydronium ion ( $\text{H}_3\text{O}^+$ ) as a proton donor in this model of QM/MM was also investigated. As **Figure 8.6J** shows, hydronium ion simultaneously causes protonation of oxygen amide and  $\text{C}_{87}\text{S}-\text{S}_{\text{C}440}$  bond formation. It is worth mentioning that high reactivity of hydronium ion and high proton affinity of the target system prevent examining the optimized system including the hydronium ion separately, but its proton donating process.

Comparing the Mn-coordination in the two optimized complexes, one a result of hydronium ion presence (**Figure 8.6J**) and the other a result of proton transfer from the hydroxyl group of Thr438 to oxygen amide (Complex **I**), indicates that hydronium ion leads to a weaker interaction of  $\text{wO}\cdots\text{Mn}$  and  $\text{E}_{411}\text{O}\cdots\text{Mn}$  with a distance of 2.47 and 2.30 Å, respectively. This is while it causes a stronger non-covalent interaction of  $\text{C}_{440}\text{S}\cdots\text{N}_{\text{D}411}$ ; 2.27 Å vs. 2.84 Å. However,  $r(\text{C}_{87}\text{S}-\text{S}_{\text{C}440})$  has not changed considerably in two complexes.

On the other hand, to have a real model of the studied system, hydronium ion was also applied to neutralize and stabilize the oxygen amide bond of Ala439 (see **Figure 8.6K**).





**Figure 8.7L** shows the optimized enzyme active site with the larger region of QM layer in which Glu411 through carboxylate group, His369, Cys87, and Asp441 are coordinated to the Mn(II) to form a penta-coordinated metal center. Comparing this model with the smaller model of the QM region does not display significant changes in metal coordination distances; all ligands are situated at 2.02-2.55 Å from Mn(II) center. Additionally,  $r(\text{C}_{440}\text{S}-\text{N}_{\text{D441}})$  in sulfenylamide ring is affected negligibly relative to the smaller model (compare **Figure 8.7L** and **8.3C**).

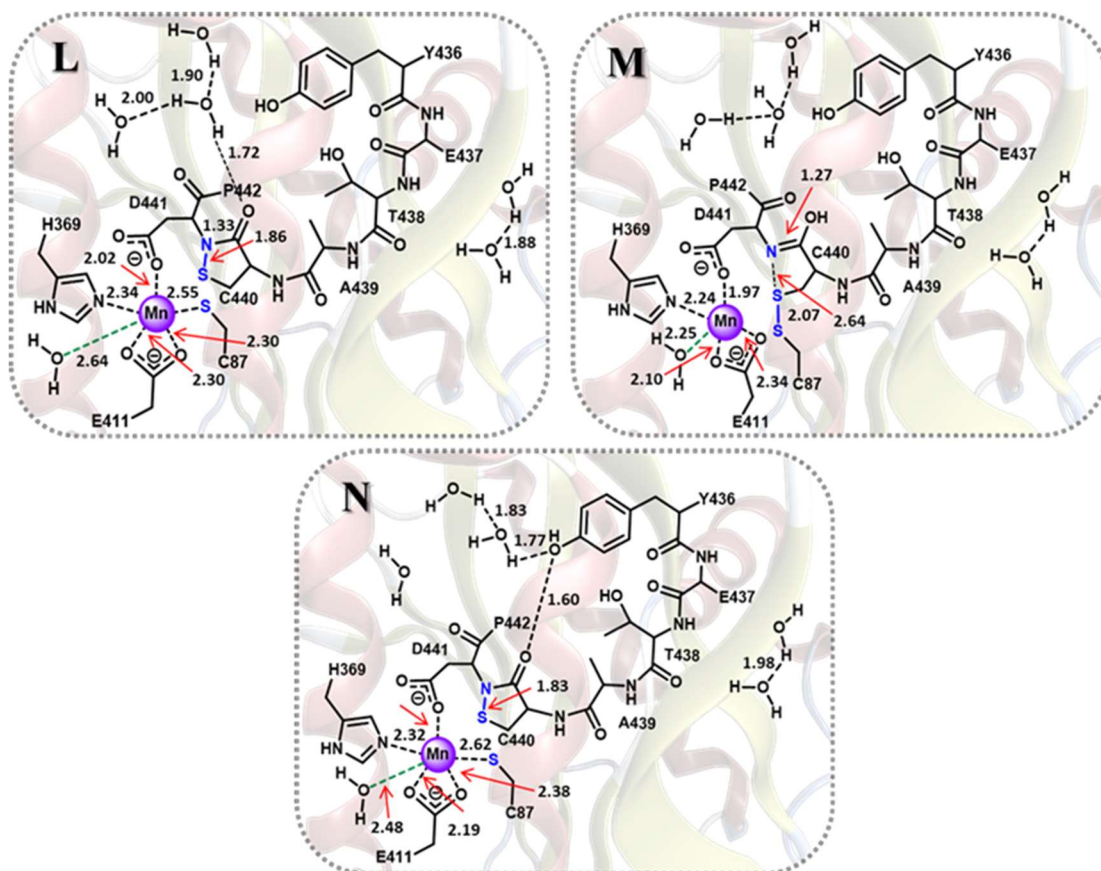
In the following, the impact of the protonated amide oxygen of cyclic sulfenylamide has been investigated. Expectedly, based on the observed results from the QM-cluster approach and QM/MM with the smaller model of QM layer, the protonated amide oxygen results in removing Cys87 thiolate from metal centered-coordination,  $\text{C}_{440}\text{S}-\text{N}_{\text{D441}}$  bond breaking, and eventually disulfide bond ( $\text{C}_{87}\text{S}-\text{S}_{\text{C440}}$ ) formation with  $\Delta E = -1752.2$  kJ/mol (see **Figure 8.7M**). The larger QM region leads to shortening the measured  $r(\text{C}_{87}\text{S}-\text{S}_{\text{C440}})$  by 0.03 to 2.07 Å comparing to the smaller QM region. Although the larger size of QM layer decreases the  $r(\text{S}\cdots\text{N})$  distance in the protonated complex by 0.62 Å from 3.26 Å in the model with the smaller size of QM region (**Figure 8.3D**) to 2.64 Å in a model with the larger size of QM region (**Figure 8.7M**).

To gain more reliable insight into the Mn(II) centered-coordination in the exchange mechanism, a water molecule has been placed in the second sphere of the metal ion in the active site which has been located at a distance of 2.64 Å from Mn(II) in the optimized complex (see **Figure 8.7L**). As observed, coincident with removing the cysteine thiolate from the metal coordination sphere to form a disulfide bond, a water molecule is replaced and ligated to the metal ion with a distance of 2.25 Å from manganese to keep the penta-ligated metal ion center (see **Figure 8.7M**). The lower calculated reaction energy ( $\Delta E = -1752.2$  kJ/mol) of disulfide-sulfenylamide exchange mechanism applying the larger QM region including water molecule compare to the smaller QM region without water molecule



in the active site ( $\Delta E = -1534.3$  kJ/mol) might imply the significant role of the water molecule in the exchange mechanism and Mn(II) coordination.

It is worth pointing out that improving the basis set applied on the sulfur atom from 6-31G(d,p) to 6-311G(d,p) results in  $r(\text{S—S})$  bond lengthening and  $r(\text{S—N})$  bond shortening in the obtained optimized complexes by  $< 0.02$  Å.



**Figure 8.7.** Schematic representation of optimized structures of the active site (L) and protonated oxygen amide in the active site (M) and a structure result of examining proton transferring from Tyr436 to oxygen atom of sulfenylamide bond (N) with the larger model of QM region using ONIOM(B3LYP/6-31G(d,p):AMBER96) with the selected distances in ångstroms (Å).

As observed previously, in both QM-cluster and QM/MM with a small model of QM region, the  $\text{C}_{440}\text{S—N}_{\text{D441}}$  and neutral Tyr436 were preferred to  $\text{C}_{87}\text{S—S}_{\text{C440}}$  and

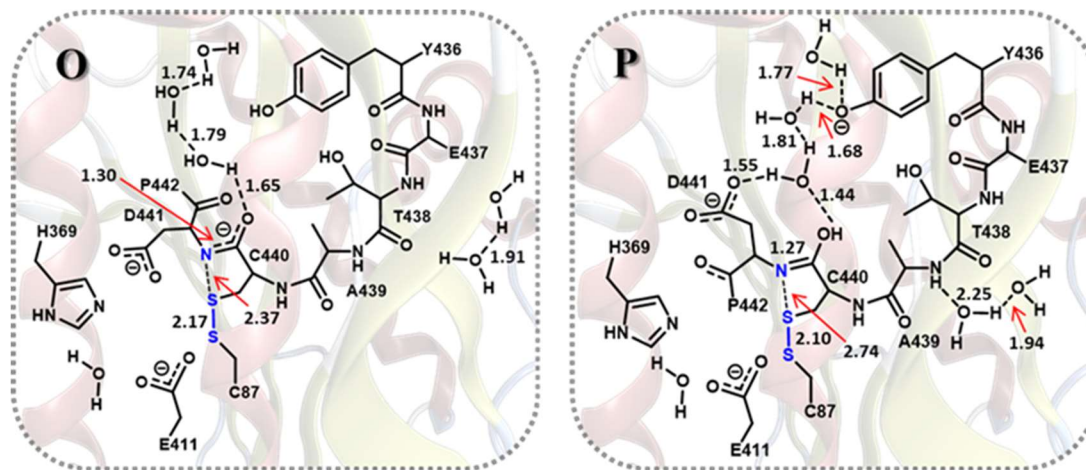
deprotonated Tyr436. In this model with the larger QM region, with the aim of stabilizing the oxyanion of the tyrosinate group through an H-bonding network, a few numbers of explicit water molecules were employed and proton transferring process from the hydroxyl group of Tyr436 to the oxygen atom of cyclic sulfenylamide, alongside with  $C_{440}S-N_{D441}$  breaking and  $C_{87}S-S_{C440}$  formation was investigated. Similarly, as shown in **Figure 8.7N**, the obtained optimized structure also favors the neutral Tyr436 and cyclic sulfenylamide formation. The obtained optimized complex is 266.7 kJ/mol lower in energy relative to the Complex shown in **Figure 8.7L**. This is while in the optimized latter complex,  $r(S\cdots Mn)$  has been increased by 0.07 Å from 2.55 to 2.62 Å. This increase is simultaneous with  $r(wO\cdots Mn)$  noticeable decrease from 2.64 to 2.48 Å (compare **Figure 8.7L** and **8.7N**).

#### 8.3.3.1 Mn(II) Influence on the Exchange Mechanism

Mn(II) contribution in the  $C_{440}S-N_{D441}$  and  $C_{87}S-S_{C440}$  switching mechanism was also examined while it was taken out of the enzyme active site. As represented in **Figure 8.8O**, the larger QM region and Mn(II) absence result in  $r(S-N)$  distance lengthening from 1.86 to 2.37 Å. This bond cleavage also leads to S—S bond forming,  $r(S-S)$ : 2.17 Å. It is worth stating that the applied larger QM region influence both sulfur-nitrogen and sulfur-sulfur interaction more considerably relative to the corresponding complex with a smaller model of QM region (compare **Figure 8.4F** and **8.8O**); 2.03 vs. 2.37 Å and 2.44 vs. 2.17 Å, the measured distances of  $r(S\cdots N)$  and  $r(S\cdots S)$  using the small and large model of QM region, respectively.

Additionally, in the absence of Mn(II) in the enzyme active site, Tyr436's ability to act as a proton donor to the oxygen atom of cyclic sulfenylamide, coincident with S—S bond formation was also studied. In contrast with the previously studied models (QM-cluster and QM/MM) including the metal ion in the active site, S—S bond formation is favored over S—N in the metal ion absence (**Figure 8.8P**). Comparison with the corresponding complex with the smaller QM region indicates that large QM region has affected the  $S\cdots N$

interaction more significantly than  $S\cdots S$  interaction; a 0.23 Å increase in  $r(S\cdots N)$  vs. 0.08 Å decrease in  $r(S\cdots S)$ ; Compare **Figure 8.8P** and **8.4G**.



**Figure 8.8.** Schematic illustration of optimized complexes of the active site (**O**) and the result of proton transferred from Tyr436 to oxygen amide in the active site (**P**) with the larger model of QM region in the absence of metal ion with selected distances shown in ångstroms (Å) using ONIOM(B3LYP/6-31G(d,p):AMBER96).

Furthermore, applying the larger QM region involving a few numbers of explicit water molecules and their contribution in the formation of strong H-bonding networks simultaneous with proton transferring from Tyr436 to the amide oxygen atom leads to the formation of a complex as shown in **Figure 8.8P** that lies 1011.5 kJ/mol down in energy with respect to the represented optimized complex in **Figure 8.8O** with the neutral Tyr436.

### 8.3.4 Molecular Dynamics (MD) Simulations

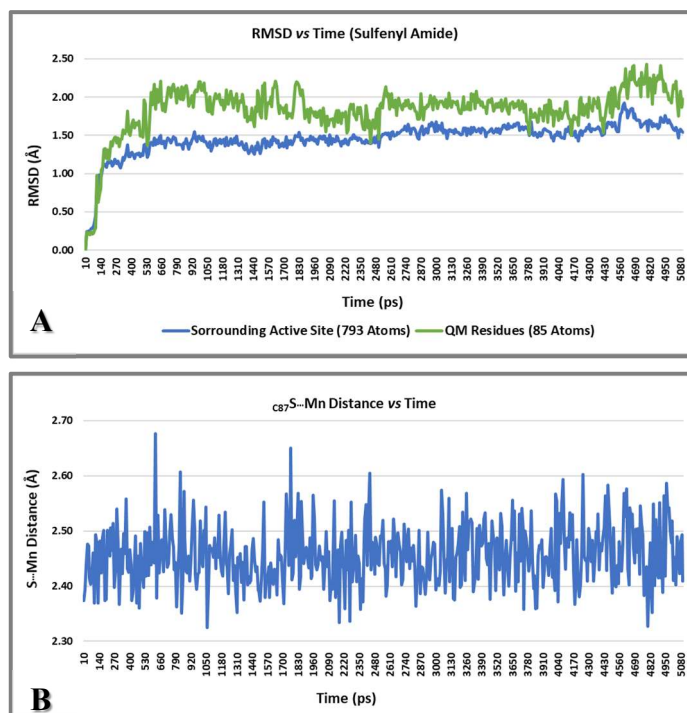
To evaluate the structural consequences and changes of the enzyme active site and understand the  $C_{87}S\cdots Mn$  distance, also the enzymatic conformational behavior during the exchange mechanism while Cys87 has formed a disulfide bond ( $C_{87}S-S_{C440}$ ) and has been removed from Mn(II) first coordination shell, a series of MD simulations has been

performed. As mentioned in the computational methodology section, 100 ps equilibration simulation was followed by a 5 ns production run to investigate the active site consistency.

First, the enzyme with the cyclic sulfenylamide species was examined. **Figure 8.9** represents the RMSD of the QM layer including Mn(II), His369, Asp441, Glu411, Cys87 and 440, and Tyr436 (in total 85 atoms). As shown, the RMSD of the generated structures fluctuates with a range of  $1.86 \pm 0.23$  Å during a 5 ns simulation. Additionally, the RMSD of a larger region of the system including the QM layer residues and their surrounding (a total of 793 atoms) was plotted in **Figure 8.9**. Obviously, the lower fluctuation range ( $1.48 \pm 0.15$  Å) was observed relative to the previous RMSD indicating the stability of the system.

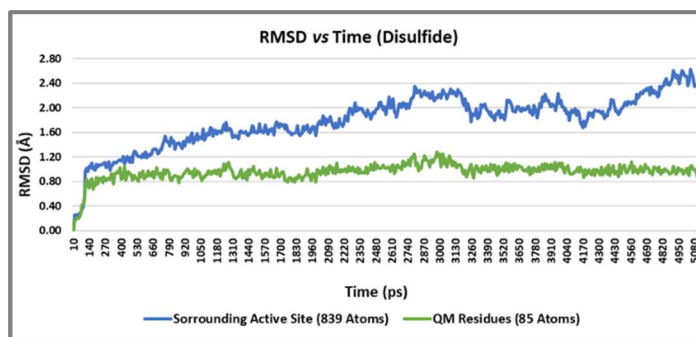
As one of the main criteria in molecular dynamics studies, the fluctuation of manganese and sulfur distance ( $C_{87}S \cdots Mn$ ) has been investigated during the simulation. As it is obvious (**Figure 8.9**), this distance approximately fluctuates with a range of  $2.46 \pm 0.05$  Å. The mean distance also compromises well with obtained distance using quantum mechanical methods.

All these results show that the formation and presence of the sulfenylamide species containing  $C_{440}S-N_{D441}$  and  $C_{87}S \cdots Mn$  coordination in the active site of this enzyme is feasible despite the longer observed distance of  $C_{87}S \cdots Mn$  in the crystal structure (2.82 Å).



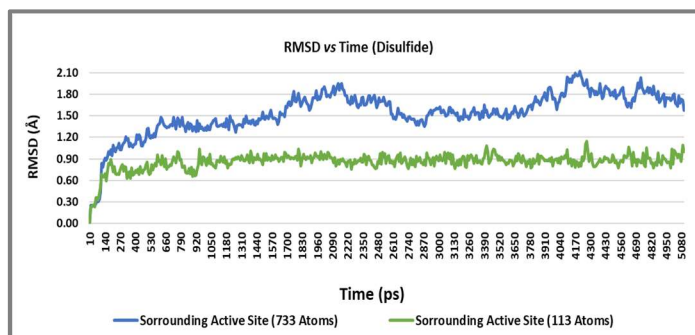
**Figure 8.9.** Plot of the calculated RMSD of QM layer and its surrounding (**A**) and distance change between  $c_{87}S$  and Mn(II) with respect to the time (**B**) in a system including cyclic sulfenylamide.

Furthermore, the enzyme stability during the exchange mechanism, while a proton from Tyr436 has been transferred to amide oxygen of cyclic sulfenylamide and  $c_{440}S-S_{C87}$  has been formed, was also studied. The RMSD plot of the QM region (Mn(II), His369, Asp441, Glu411, Cys87 and 440, and Tyr436) with a total number of 85 atoms is shown in **Figure 8.10**. The generated structures show a fluctuation of  $0.97 \pm 0.09$  Å after equilibrium stage (100 ps) during a 5 ns production run. Although a bigger range of fluctuation,  $1.82 \pm 0.37$  Å, has been observed while a larger region of the system including QM residues and neighboring (839 atoms in total) were selected (see **Figure 8.10**) indicating the less stability of the system with the disulfide bond formation.



**Figure 8.10.** RMSD of the QM layer and its surrounding in a system including disulfide.

As previously mentioned and observed in the crystal structure, the second active site in the enzyme involves disulfide rather than cyclic sulfenylamide species. The calculated RMSDs of this enzyme active site with two different-size models of active site residues and their surroundings are presented in **Figure 8.11**. The calculated RMSD of the active site (including 113 atoms) shows a fluctuation of  $0.87 \pm 0.08$  Å while the increasing the number of residues in the active site to 733 atoms raises the fluctuation to a range of  $1.57 \pm 0.24$  Å. It is worth stating that the observed fluctuations are less than the corresponding ones in the first active site with disulfide (compare **Figure 8.10** and **8.11**).



**Figure 8.11.** Plot of the second active site and its surrounding RMSDs in the enzyme including disulfide.

## 8.4 Conclusion

Regarding the rarity and importance of the observed cyclic sulfenylamide moiety including S—N in the active site of DAH7PS enzyme and its capability to be involved in a switching

mechanism to form a disulfide (S—S), the present study has investigated this manganese metalloenzyme active site and the influential environmental factors which have potentially impacted this competitive reaction between S—S and S—N formations applying a multi-scale computational approach.

Our initial findings from the QM-cluster only approach present the preference of the cyclic sulfenylamide ( $C_{440}S-N_{D441}$ ) to disulfide ( $C_{87}S-S_{C440}$ ) in the enzyme active site, although applying the acidic environment leads to removing Cys87 from Mn(II)-coordinated center and preference of disulfide formation.

With these initial obtained observations, we extended our studies into ONIOM(QM/MM) calculations to gain more accurate and reliable findings, while the protein environment with its strict and long-range interactions are taken into account.

Applying the ONIOM(QM/MM) approach with only the key residues in the enzyme active site in the QM region (total of 83 atoms) also confirmed the obtained results from QM-cluster approach: S—N preference to S—S in the active site *vs.* S—S preference to S—N in the acidic environment. The only exception has been observed in the Mn(II) coordination center, in which removing the Cys87 from Mn(II)-coordinated center to form the disulfide has not been replaced with any ligations, but stronger interactions with the remained ligated residues results from the protein environment.

Influence of the metal ion (Mn(II)) in the exchange mechanism has also been investigated. The results imply the significant contribution of the Mn(II) to studied competitive reaction. Excluding the metal ion from the enzyme active site leads to lengthening the  $r(S-N)$  distance *vs.* shortening the  $r(S\cdots S)$  comparison with the active site including Mn(II) using B3LYP. Notably, applying the empirical D3 dispersion correction (B3LYP-D3) confirms the formation of the strong covalent interaction of S—S in the absence of metal ion: S—S preference to S—N in the Mn(II)-excluded active site.

Also, the absence of metal ion in the enzyme active site causes Tyr436 to act as a proton donor and formation of disulfide. It is while the presence of Mn(II) in the enzyme active site hinders the Tyr436 to behave as a Lewis acid (confirmed by QM-cluster and QM/MM).

Furthermore, gaining more accurate and detailed insight into the local surrounding of the enzyme active site and metal ion coordination, the ONIOM(QM/MM) study was also expanded to the model with a larger size of QM region (in total 127 atoms). Regarding the switching disulfide-sulfenylamide mechanism, the obtained results agree with the model using the smaller size of the QM layer. However, applying the larger size of the QM region causes the stronger interaction of  $S \cdots N$  in the acidic environment vs. corresponding interaction using the smaller size of the QM region.

With the aim of the study Tyr436 ability to behave as a Lewis acid while a few numbers of explicit water molecules were employed in the QM layer to stabilize the oxyanion of the tyrosinate group throughout the expanded H-bonding network (in contrast with the small model of QM layer and QM-cluster approach) and S—S formation, the results once more imply the S—N preference to S—S and a neutral Tyr436 in the enzyme active site.

Importantly, coincident with disulfide-sulfenylamide exchange mechanism, Mn(II) coordination has also been explored. Comparing the metal ion ligations in both applied models of QM layers indicate that Mn(II) is coordinated to Glu411 (through carboxylate group), His369, Cys87, and Asp441 to form a penta-coordinated metal center. It is while removing the cysteine thiolate (Cys87) from the metal coordination sphere to form a disulfide bond, the remaining open coordination is occupied with a water molecule to keep the penta-coordinated geometry.

As well as, taking out the metal ion from the enzyme active site to accurately examine the Mn(II) importance in the exchange mechanism indicates the preference of the covalent S—S to S—N formation. The obtained results agree with observed findings applying the empirical D3 dispersion correction (B3LYP-D3) with the small size of the QM layer. Similarly, it has been observed that the absence of the metal ion in the enzyme active site



results in Tyr436 to act as a Lewis acid and finally disulfide formation. Although, applying the larger size of the QM region affects S $\cdots$ N interaction more significantly than S $\cdots$ S interaction, relative to the corresponding ones using the smaller size of the QM layer.

Also, in this present study, the contribution of Thr438 positioned at the neighboring cyclic sulfenylamide in the switching mechanism and presence of the hydronium ion in the enzyme active site have been investigated. All the results indicate providing an acidic environment leads Cys87 to remove metal ion coordination to form a disulfide bond which is simultaneous with sulfur-nitrogen bond cleavage.

### 8.5 References

1. Harmel, R.; Fiedler, D., Features and Regulation of Non-Enzymatic Post-Translational Modifications. *Nat. Chem. Biol.* **2018**, *14*, 244.
2. Walsh, G., Post-translational Modifications of Protein Biopharmaceuticals. *Drug Discov. Today* **2010**, *15*, 773-780.
3. Carrico, I. S., Chemoselective Modification of Proteins: Hitting the Target. *Chem. Soc. Rev.* **2008**, *37*, 1423-1431.
4. Walsh, C. T.; Garneau-Tsodikova, S.; Gatto Jr., G. J., Protein Posttranslational Modifications: The Chemistry of Proteome Diversifications. *Angew. Chem. Int. Ed.* **2005**, *44*, 7342-7372.
5. Duan, G. Y.; Walther, D., The Roles of Post-translational Modifications in the Context of Protein Interaction Networks. *Plos Comput. Biol.* **2015**, *11*, 1-23.
6. Knorre, D. G.; Kudryashova, N. V.; Godovikova, T. S., Chemical and Functional Aspects of Posttranslational Modification of Proteins. *Acta Naturae* **2009**, *1*, 29-51.
7. Green, K. D.; Garneau-Tsodikova, S., 5.15 - Posttranslational Modification of Proteins. In *Comprehensive Natural Products II*, Liu, H.-W.; Mander, L., Eds. Elsevier: Oxford, 2010; pp 433-468.

8. Sarge, K. D.; Park-Sarge, O.-K., Chapter Four - SUMO and Its Role in Human Diseases. In *International Review of Cell and Molecular Biology*, Jeon, K. W., Ed. Academic Press: 2011; Vol. 288, pp 167-183.
9. Dhuriya, Y. K.; Sharma, D.; Naik, A. A., Cellular demolition: Proteins as Molecular Players of Programmed Cell Death. *Int. J. Biol. Macromol.* **2019**.
10. Fass, D.; Thorpe, C., Chemistry and Enzymology of Disulfide Cross-Linking in Proteins. *Chem. Rev.* **2018**, *118*, 1169-1198.
11. Maulik, V. T.; Jennifer, S. L.; Teruna, J. S., The Role of Thiols and Disulfides on Protein Stability. *Current Protein Pept. Sci.* **2009**, *10*, 614-625.
12. Tichá, T.; Lochman, J.; Činčalová, L.; Luhová, L.; Petřivalský, M., Redox Regulation of Plant S-nitrosoglutathione Reductase Activity Through Post-Translational Modifications of Cysteine Residues. *Biochem. Bioph. Res. Co.* **2017**, *494*, 27-33.
13. Kim, H.-J.; Ha, S.; Lee, H. Y.; Lee, K.-J., ROSics: Chemistry and Proteomics of Cysteine Modifications in Redox Biology. *Mass Spectrom. Rev.* **2015**, *34*, 184-208.
14. Iciek, M.; Kowalczyk-Pachel, D.; Bilaska-Wilkosz, A.; Kwiecień, I.; Górny, M.; Włodek, L., S-sulfhydration as a Cellular Redox Regulation. *Biosci. Rep.* **2016**, *36*, e00304.
15. Couturier, J.; Chibani, K.; Jacquot, J.-P.; Rouhier, N., Cysteine-Based Redox Regulation and Signaling in Plants. *Frant. Plant. Sci* **2013**, *4*.
16. Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G., Chemical Modification of Proteins at Cysteine: Opportunities in Chemistry and Biology. *Chem-Asian J.* **2009**, *4*, 630-640.
17. Chung, H. S.; Wang, S.-B.; Venkatraman, V.; Murray, C. I.; Eyk, J. E. V., Cysteine Oxidative Posttranslational Modifications. *Circ. Res.* **2013**, *112*, 382-392.
18. Petkowski, J. J.; Bains, W.; Seager, S., Natural Products Containing a Nitrogen-Sulfur Bond. *J. Nat. Prod.* **2018**, *81*, 423-446.
19. Parvez, S.; Long, M. J. C.; Poganik, J. R.; Aye, Y., Redox Signaling by Reactive Electrophiles and Oxidants. *Chem. Rev.* **2018**, *118*, 8798-8888.

20. van Montfort, R. L. M.; Congreve, M.; Tisi, D.; Carr, R.; Jhoti, H., Oxidation State of the Active-Site Cysteine in Protein Tyrosine Phosphatase 1B. *Nature* **2003**, *423*, 773.
21. Sivaramakrishnan, S.; Cummings, A. H.; Gates, K. S., Protection of a Single-Cysteine Redox Switch from Oxidative Destruction: On the Functional Role of Sulfenyl Amide Formation in the Redox-Regulated Enzyme PTP1B. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 444-447.
22. Sarma, B. K.; Mugesh, G., Redox Regulation of Protein Tyrosine Phosphatase 1B (PTP1B): A Biomimetic Study on the Unexpected Formation of a Sulfenyl Amide Intermediate. *J. Am. Chem. Soc.* **2007**, *129*, 8872-8881.
23. Defelipe, L. A.; Lanzarotti, E.; Gauto, D.; Marti, M. A.; Turjanski, A. G., Protein Topology Determines Cysteine Oxidation Fate: The Case of Sulfenyl Amide Formation among Protein Families. *Plos Comput. Biol.* **2015**, *11*, 1-25.
24. Salmeen, A.; Andersen, J. N.; Myers, M. P.; Meng, T. C.; Hinks, J. A.; Tonks, N. K.; Barford, D., Redox Regulation of Protein Tyrosine Phosphatase 1B Involves a sulphenyl-Amide Intermediate. *Nature* **2003**, *423*, 769-773.
25. Ruddraraju, K. V.; Zhang, Z. Y., Covalent Inhibition of Protein Tyrosine Phosphatases. *Mol. Biosyst.* **2017**, *13*, 1257-1279.
26. Bence, K. K.; Delibegovic, M.; Xue, B.; Gorgun, C. Z.; Hotamisligil, G. S.; Neel, B. G.; Kahn, B. B., Corrigendum: Neuronal PTP1B Regulates Body Weight, Adiposity and Leptin Action. *Nat. Med.* **2010**, *16*, 237.
27. Kyriakou, E.; Schmidt, S.; Dodd, G. T.; Pfuhlmann, K.; Simonds, S. E.; Lenhart, D.; Geerlof, A.; Schrieffer, S. C.; De Angelis, M.; Schramm, K.-W.; Plettenburg, O.; Cowley, M. A.; Tiganis, T.; Tschöp, M. H.; Pfluger, P. T.; Sattler, M.; Messias, A. C., Celastrol Promotes Weight Loss in Diet-Induced Obesity by Inhibiting the Protein Tyrosine Phosphatases PTP1B and TCPTP in the Hypothalamus. *J. Med. Chem.* **2018**, *61*, 11144-11157.

28. Ezzat, S. M.; Bishbishy, M. H. E.; Habtemariam, S.; Salehi, B.; Sharifi-Rad, M.; Martins, N.; Sharifi-Rad, J., Looking at Marine-Derived Bioactive Molecules as Upcoming Anti-Diabetic Agents: A Special Emphasis on PTP1B Inhibitors. *Molecules* **2018**, *23*, 3334.
29. Dokainish, H. M.; Gauld, J. W., Formation of a Stable Iminol Intermediate in the Redox Regulation Mechanism of Protein Tyrosine Phosphatase 1B (PTP1B). *ACS Catal.* **2015**, *5*, 2195-2202.
30. Roos, G.; Messens, J., Protein Sulfenic Acid Formation: From Cellular Damage to Redox Regulation. *Free Radic. Biol. Med.* **2011**, *51*, 314-326.
31. Cremers, C. M.; Jakob, U., Oxidant Sensing by Reversible Disulfide Bond Formation. *J. Biol. Chem.* **2013**, *288*, 26489-26496.
32. Webby, C. J.; Jiao, W. T.; Hutton, R. D.; Blackmore, N. J.; Baker, H. M.; Baker, E. N.; Jameson, G. B.; Parker, E. J., Synergistic Allostery, a Sophisticated Regulatory Network for the Control of Aromatic Amino Acid Biosynthesis in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **2010**, *285*, 30567-30576.
33. Reichau, S.; Blackmore, N. J.; Jiao, W. T.; Parker, E. J., Probing the Sophisticated Synergistic Allosteric Regulation of Aromatic Amino Acid Biosynthesis in *Mycobacterium tuberculosis* Using.- Amino Acids. *Plos One* **2016**, *11*, 1-18.
34. Blackmore, N. J.; Nazmi, A. R.; Hutton, R. D.; Webby, M. N.; Baker, E. N.; Jameson, G. B.; Parker, E. J., Complex Formation between Two Biosynthetic Enzymes Modifies the Allosteric Regulatory Properties of Both an Example of Molecular Symbiosis. *J. Biol. Chem.* **2015**, *290*, 18187-18198.
35. Zhou, L.; Wu, J.; Janakiraman, V.; Shumilin, I. A.; Bauerle, R.; Kretsinger, R. H.; Woodard, R. W., Structure and Characterization of the 3-Deoxy-d-arabino-heptulosonate 7-phosphate Synthase from *Aeropyrum pernix*. *Bioorg. Chem.* **2012**, *40*, 79-86.
36. *Molecular Operating Environment (MOE), 2018.01; Chemical Computing Group ULC.*, 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2019.

37. Neves, R. P. P.; Sousa, S. F.; Fernandes, P. A.; Ramos, M. J., Parameters for Molecular Dynamics Simulations of Manganese-Containing Metalloproteins. *J. Chem. Theory Comput.* **2013**, *9*, 2718-2732.
38. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K., Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781-1802.
39. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; N. Rega; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; J. V. Ortiz; Cioslowski, J.; Fox, D. J. *Gaussian 09*, E.01; Gaussian, Inc.: Wallingford CT, 2015.
40. Blomberg, M. R. A.; Borowski, T.; Himo, F.; Liao, R.-Z.; Siegbahn, P. E. M., Quantum Chemical Studies of Mechanisms for Metalloenzymes. *Chem. Rev.* **2014**, *114*, 3601-3658.
41. Himo, F., Recent Trends in Quantum Chemical Modeling of Enzymatic Reactions. *J. Am. Chem. Soc.* **2017**, *139*, 6780-6786.
42. Lind, M. E. S.; Himo, F., Quantum Chemical Modeling of Enantioconvergency in Soluble Epoxide Hydrolase. *ACS Catal.* **2016**, *6*, 8145-8155.

43. Lee, C. T.; Yang, W. T.; Parr, R. G., Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. *Phys. Rev. B* **1988**, *37*, 785-789.
44. Becke, A. D., A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372-1377.
45. Sheng, X.; Zhu, W.; Huddleston, J.; Xiang, D. F.; Raushel, F. M.; Richards, N. G. J.; Himo, F., A Combined Experimental-Theoretical Study of the LigW-Catalyzed Decarboxylation of 5-Carboxyvanillate in the Metabolic Pathway for Lignin Degradation. *ACS Catal.* **2017**, *7*, 4968-4974.
46. Prejanò, M.; Marino, T.; Russo, N., QM Cluster or QM/MM in Computational Enzymology: The Test Case of LigW-Decarboxylase. *Front. Chem.* **2018**, *6*, 1-9.
47. Yamanaka, S.; Kanda, K.; Saito, T.; Kitagawa, Y.; Kawakami, T.; Ehara, M.; Okumura, M.; Nakamura, H.; Yamaguchi, K., Does B3LYP Correctly Describe Magnetism of Manganese Complexes with Various Oxidation Numbers and Various Structural Motifs? *Chem. Phys. Lett.* **2012**, *519-520*, 134-140.
48. Gennari, M.; Brazzolotto, D.; Yu, S.; Pécaut, J.; Philouze, C.; Rouzières, M.; Clérac, R.; Orio, M.; Duboc, C., Effect of the Metal on Disulfide/Thiolate Interconversion: Manganese versus Cobalt. *Chem-Eur. J.* **2015**, *21*, 18770-18778.
49. Grimme, S., Density Functional Theory with London Dispersion Corrections. *WIREs. Comput. Mol. Sci.* **2011**, *1*, 211-228.

## **CHAPTER 9**

### **Conclusions and Future Work**



## 9.1 Conclusions

In this work, multi-scale computational modeling techniques have been applied to explore the sulfur chemistry, its structure, functionality, and reactivity in a wide variety of sulfur-containing systems.

In **chapter 3**, the ability of the DFT functionals (B3LYP, B3PW91,  $\omega$ B97XD, M06-2X, and M08-HX) relative to electron correlation methods MP2 and QCISD to produce reliable and accurate structures, as well as thermochemical data for sulfur/selenium-containing systems, was investigated. Bond lengths, proton affinities (PA), gas phase basicities (GPB), chalcogen–chalcogen bond dissociation enthalpies (BDE), and the hydrogen affinities (HA) of thiyl/selenyl radicals were evaluated for a range of  $RX_n(H)$  ( $X=S, Se, R=CH_3, CH_2CH_3$ , and cysteine,  $n=1-4$ ). The S–S bond length was found to be the most sensitive to basis set choice, while the geometry of selenium-containing compounds was less sensitive to the basis set. In mixed chalcogens species of sulfur and selenium, the location of the sulfur atom affects the S–Se bond length as it can hold more negative charge. PA, GPB, BDE, and HA of selenium systems were all lower, indicating more acidity and more stability of radicals. Extending the sulfur chain in cysteine results in a decrease of BDE and HA, but these plateau at a certain point ( $199 \text{ kJ mol}^{-1}$  and  $295 \text{ kJ mol}^{-1}$ ), and PA and GPB are also decreased relative to the thiol, indicating that the polysulfur species exist as thiolates in a biological system. In general, it was found that  $\omega$ B97XD/6-311G(2d,p) gave the most reasonable structures and thermochemistry relative to benchmark calculations.

In **chapter 4**, applying two different computational solvation models, implicit and hybrid implicit/explicit solvation techniques, the influence of individual solvent molecules, solvent-solute interactions, and in general reaction environment and its constituents on the mechanism of sulfur-nitrogen bond formation were discussed. To this goal, formation of the sulfonamide ( $RS(O)_2-NR'R''$ ) as one of the precious and versatile highly functionalized sulfur species from N-tosyl hydrazones, amines, and sulfur dioxide in a wide variety of solvents with different dielectric constants was studied. Our findings reveal the significant



contribution of solvent-solute interactions in the reaction progress. The studied ene reaction between the sulfur dioxide and N-tosyl hydrazine toward the  $\beta,\gamma$ -unsaturated sulfinic acid formation shows how the applied explicit solvent molecules with the higher proton affinity lead to a noticeable decrease in the reaction energy, while the changes in activation energy are less significant. Hence, the study introduces dimethyl sulfide (DMSO) as a suitable solvent and base for the mentioned ene reaction. The reactivity of sulfene a result from the decomposition of the  $\beta,\gamma$ -unsaturated sulfinic acid toward amine (piperidine in the present study) to form sulfonamide as the desired end product was explored. As results indicate the presence of base with its active contribution *via* formation a 6-cyclic transition structure (*e.g.*, water) in proton transferring process from the nitrogen atom of amine to the carbon atom of sulfene plays a considerable role in decreasing the activation barrier. However, presence of a base with lower proton affinity than amine (*e.g.*, DMSO) causes an increase in the activation barrier. Consequently, employing the excess amount of piperidine acting as both a nucleophile to form a sulfur-nitrogen bond and a base to assist proton transferring process along the addition reaction leads to a notable decrease in the activation barrier (by  $\sim 7$  times).

**In chapter 5**, we explored the formation of highly efficient  $\text{SO}_2$ -containing molecules including sulfones, sulfonamides, and sulfamides *via* a radical-based reaction from two equivalents of a systematic series of xiamycin-inspired aromatic C- and N-centered radicals and sulfur dioxide, in the gas-phase and aqueous solution. Using the *ab initio* MP2 method and DFT functionals, the formation of such compounds was extensively discussed. Our investigation shows that C–S(O<sub>2</sub>) bond formation is favoured over N–S(O<sub>2</sub>) which necessarily does not agree with the observed spin density distribution in the initial heterocyclic radical species. Consequently, it results in the introduction of sulfones ( $\text{R}_3\text{C-SO}_2\text{-CR}_3$ ) and sulfamides ( $\text{R}_2\text{N-SO}_2\text{-NR}_2$ ) species as the most and least thermodynamically preferred  $\text{SO}_2$ -containing molecules in the present studies, respectively. Importantly, the observed results of the formation of relevant  $\text{SO}_2$ -containing molecules from radical

xiamycin are in a good agreement with experimentally observed diaryl sulfonamide and sulfone antibiotics which were identified as bioproducts through a bacterial synthesis of flavoenzyme *XiaH* and sulfur dioxide capturing. The ability of some DFT functionals (B3LYP, B3LYP-D3, and M06-2X) and conventional *ab initio* methods were assessed in this study. The results represent the successful performance of M06-2X as the *meta*-GGA functional in predicting the geometrical parameters and energies in studying relevant SO<sub>2</sub>-containing molecules.

In **Chapter 6**, using DFT-based methods reactivity of highly electrophilic HNO toward cysteinyl and cysteinyl persulfide, as one of the important post-translationally modified cysteinyl with enhanced nucleophilicity was investigated. The results indicate that formation of Cys-X-NHOH (X=S and S-S), as the first and chief intermediate is independent of the peptide's position in the amino acid chain. Although the potential energy surface for the formation of such S–N containing species are potentially affected by the peptide's position and polarity of the biological environment. Significantly, our findings show that further modifications of Cys-X-NHOH (X=S and S-S) intermediate result in a wide variety of sulfur-containing biomolecules dependence upon the peptide's position which was initially derived from. Cys-X-NH<sub>2</sub><sup>+</sup>OH (X=S and S-S) derived from the internal peptide leads to the formation of Cys-X(O)-NH<sub>2</sub> (X=S and S-S), and Cys-XX-Cys (X=S) *via* a rearrangement reaction and a nucleophilic substitution, respectively. It is while Cys-X-NH<sub>2</sub><sup>+</sup>OH (X=S and S-S) derived from the C-terminus peptides potentially affected by the presence of carboxylic acid and results in the formation of Cys-X-OH (X=S, S-S) through the intermediacy of a 5- or 6-membered cyclic structure in cysteinyl and cysteinyl persulfide, respectively and independent of reactive oxygen species (ROS).

In **chapter 7**, we used multi-scale computational methods to provide significant insights into the detailed understanding of the bacterial Ni-DddK-catalyzed dimethylsulfoniopropionate (DMSP) cleavage, as a sulfur-containing compound to form dimethyl sulfide (DMS). Applying the QM-cluster only approach and quantum

mechanics/molecular mechanics (QM/MM) techniques, firstly we investigated the bound and unbound active sites to gain atomistic explanations of the key Ni-coordinated ligands and their contributions in the active site. Our results demonstrate that DMSP binding displaces the water and Tyr64 in the active site. The detachment of the Ni-coordinated Tyr64 and its removal from the first metal coordination sphere causes it directly to act as a Lewis base to initiate the  $\beta$ -concerted elimination mechanism to toward DMSP cleavage. In addition, the effects of substituting Ni(II) with Mn(II) on the catalytic cleavage mechanism has been examined. Findings imply the importance of transition metal ions acting as the cofactor in the Dddk enzyme to facilitate the DMSP lyase and formation of DMS. Two different *meta*-GGA functionals (M06 and M06L) were considered in this study. Despite the observed close agreement, the M06L functional (0%HF) exchange contribution gives a better description of this transition metal ion-containing enzyme.

In **chapter 8**, we discussed the first recognized occurrence of disulfide–sulfenylamide shuttling reaction in the active site of DAH7PS from *Mycobacterium tuberculosis*. The enzyme active site investigation showed that the identified cyclic sulfenylamide result of the thiolate Cys440 and amide nitrogen of Asp441 ( $C_{440}S-N_{D441}$ ) can be potentially involved in the switching mechanism to form a disulfide ( $C_{87}S-S_{C440}$ ). Hence, multi-scale computational modeling methods were used to explore such a process. Applying the QM-cluster approach, the first insight into the switching mechanism was obtained: preference of cyclic sulfenylamide to disulfide in the enzyme active site, while Cys87 is ligated to the metal center (Mn(II)). Afterward, employing ONIOM(QM/MM) gave a deeper insight into the mechanism, influential environmental features, and metal ion coordination in the active site. To this goal, we systematically increased the size of the QM layer to find each residue's contribution to the switching mechanism. Our results reveal the key role of Mn(II) in the potential interconversion between the cyclic sulfenylamide and disulfide conformations; preference of N–S to S–S in the active site while metal ion absence makes Tyr436 act as a Lewis acid, removing Cys87 from the metal coordination center, and finally

S–S formation. Significantly, the study shows providing an acidic environment leads to the preference of disulfide to sulfenylamide in the enzyme active site.

Everyday many novel highly sulfur-functionalized containing components, biomolecules, and natural products, specifically those including S–X (S= S, N, O), are biosynthesized, identified, and isolated from a variety of sources with unique potentials and properties, and important biological, pharmacological, and physiological functions and activities. With the fast-growing number of identified biosystems and biocatalysts being evolved in constructions such sulfur-containing species especially those including the oxidized sulfur, reactive sulfur species (RSS), polysulfur, and sulfur-nitrogen bond further computational elucidations and detailed investigations are critically needed. Unfortunately, the chemistry of their occurrence, formation, and functionalities are still challenging in experimental chemistry and have lagged behind others. Furthermore, many rare and unusual post-translational modifications of sulfur-containing peptides and proteins have brought attentions to biosynthesis and biological chemists in recent years. The present results provide experimental and computational chemists key insights into potential pathways for the synthesis of new sulfur-containing species. Furthermore, we have developed and characterized computational methodologies that are comparatively cheap yet able to provide reliable and accurate structures and thermochemical values of such species. Such knowledge is crucial for future computational studies on these and related species. That is, we have provided the essential foundation upon which future related computational studies can now build, confident in the accuracy and predictive abilities of their results.

In addition, we have deconstructed and analyzed some of the fundamental factors that must be considered when experimentally synthesizing sulfonamides and other species, including, the choice of solvent and catalytic base. Our results thus will help experimental chemists to now systematically design reaction conditions that can enable the design and synthesis of related as well as new sulfur-containing species (*e.g.*, sulfonamides). Coupled

with our new insights into the mechanisms of several physiologically key enzymes, this will enable the design of new therapeutic drugs and materials. In addition, our proof-of-concept/principle results regarding the possible interconversion of cystine and sulfenylamides in enzymes will now enable researchers to identify such mechanisms in other enzymes. In fact, it provides a brand new mechanism for enzyme regulation. As a result, it can potentially lead to the development of new therapeutic drugs or even bioengineered catalysts.

## **APPENDICES**

### **Appendix A**

**Appendix A3.1.** Selected optimized bond lengths in ångström (Å) for CH<sub>3</sub>XXH and CH<sub>3</sub>XX<sup>-</sup> (X=S, Se) using MP2.

Method	Basis set	CH <sub>3</sub> SSH			CH <sub>3</sub> SS <sup>-</sup>		CH <sub>3</sub> SeSeH			CH <sub>3</sub> SeSe <sup>-</sup>	
		C-S	S-S	S-H	C-S	S-S	C-Se	Se-Se	Se-H	C-Se	Se-Se
MP2	6-31G(d)	1.812	2.062	1.346	1.812	2.089	1.962	2.328	1.489	1.965	2.353
	6-311G(d)	1.808	2.073	1.346	1.807	2.104	1.961	2.348	1.483	1.965	2.374
	6-311G(d,p)	1.805	2.073	1.338	1.804	2.103	1.957	2.348	1.464	1.961	2.374
	6-311+G(d,p)	1.805	2.074	1.339	1.805	2.098	1.957	2.348	1.464	1.962	2.371
	6-311G(2d,p)	1.818	2.082	1.337	1.819	2.107	1.952	2.350	1.464	1.958	2.380
	6-31G11(df,p)	1.799	2.044	1.341	1.798	2.067	1.947	2.321	1.470	1.952	2.343
	6-311+G(2df,p)	1.809	2.053	1.345	1.810	2.069	1.948	2.319	1.469	1.954	2.340
	6-311++G(3df,3pd)	1.802	2.042	1.338	1.802	2.059	1.946	2.322	1.465	1.952	2.341

**Appendix A3.2.** C–S and S–S bond lengths (ångström, Å) for CH<sub>2</sub>CHSSH and CH<sub>2</sub>CHSS<sup>−</sup>.

Method	Basis set	CH <sub>2</sub> CHSSH		CH <sub>2</sub> CHSS <sup>−</sup>	
		C–S	S–S	C–S	S–S <sup>−</sup>
<b>B3LYP</b>	6-31G(d)	1.781	2.099	1.764	2.111
	6-311G(d)	1.779	2.111	1.762	2.131
	6-311+G(d)	1.778	2.109	1.756	2.124
	6-311+G(d,p)	1.778	2.110	1.755	2.123
	6-311G(2d,p)	1.777	2.096	1.759	2.109
	6-311+G(df,p)	1.776	2.096	1.753	2.109
	6-311+G(2df,p)	1.771	2.078	1.747	2.081
	6-311++G(3df,3pd)	1.767	2.069	1.742	2.076
<b>B3PW91</b>	6-31G(d)	1.772	2.078	1.753	2.088
	6-311G(d)	1.769	2.088	1.751	2.105
	6-311+G(d)	1.768	2.086	1.745	2.098
	6-311+G(d,p)	1.768	2.088	1.745	2.098
	6-311G(2d,p)	1.767	2.074	1.747	2.084
	6-311+G(df,p)	1.766	2.075	1.743	2.084
	6-311+G(2df,p)	1.762	2.058	1.737	2.063
	6-311++G(3df,3pd)	1.758	2.050	1.733	2.054
<b>ωB97XD</b>	6-31G(d)	1.766	2.075	1.755	2.090
	6-311G(d)	1.762	2.058	1.753	2.108
	6-311+G(d)	1.758	2.050	1.748	2.102
	6-311+G(d,p)	1.766	2.075	1.748	2.101
	6-311G(2d,p)	1.762	2.058	1.750	2.088
	6-311+G(df,p)	1.769	2.072	1.746	2.088
	6-311+G(2df,p)	1.764	2.057	1.740	2.069
	6-311++G(3df,3pd)	1.760	2.049	1.736	2.061
<b>M06-2X</b>	6-31G(d)	1.773	2.075	1.754	2.092
	6-311G(d)	1.772	2.083	1.752	2.108
	6-311+G(d)	1.770	2.083	1.747	2.102
	6-311+G(d,p)	1.770	2.084	1.747	2.102
	6-311G(2d,p)	1.770	2.071	1.751	2.090
	6-311+G(df,p)	1.768	2.071	1.744	2.086
	6-311+G(2df,p)	1.764	2.056	1.739	2.069
	6-311++G(3df,3pd)	1.761	2.050	1.735	2.061
<b>M08-HX</b>	6-31G(d)	1.773	2.073	1.755	2.091
	6-311G(d)	1.771	2.082	1.753	2.106
	6-311+G(d)	1.770	2.080	1.748	2.100
	6-311+G(d,p)	1.770	2.081	1.747	2.100
	6-311G(2d,p)	1.769	2.071	1.750	2.090
	6-311+G(df,p)	1.768	2.070	1.745	2.085
	6-311+G(2df,p)	1.763	2.055	1.739	2.068
	6-311++G(3df,3pd)	1.805	2.041	1.736	2.060
<b>QCISD</b>	6-311+G(2df,p)	1.771	2.069	1.750	2.086



**Appendix A3.3.** C–Se and Se–Se bond lengths (ångström, Å) for CH<sub>2</sub>CHSeSeH and CH<sub>2</sub>CHSeSe<sup>−</sup>.

Method	Basis set	CH <sub>2</sub> CHSeSeH		CH <sub>2</sub> CHSeSe <sup>−</sup>	
		C–Se	Se–Se	C–Se	Se–Se <sup>−</sup>
<b>B3LYP</b>	6-31G(d)	1.921	2.347	1.908	2.362
	6-311G(d)	1.927	2.372	1.915	2.388
	6-311G(d,p)	1.926	2.373	1.914	2.388
	6-311+G(d,p)	1.926	2.373	1.912	2.386
	6-311G(2d,p)	1.922	2.375	1.911	2.391
	6-311G(df,p)	1.921	2.360	1.908	2.373
	6-311+G(2df,p)	1.919	2.358	1.905	2.371
	6-311++G(3df,3pd)	1.918	2.359	1.905	2.371
<b>B3PW91</b>	6-31G(d)	1.908	2.324	1.894	2.337
	6-311G(d)	1.914	2.347	1.902	2.360
	6-311G(d,p)	1.913	2.347	1.901	2.360
	6-311+G(d,p)	1.913	2.347	1.899	2.359
	6-311G(2d,p)	1.910	2.350	1.898	2.364
	6-311G(df,p)	1.908	2.334	1.894	2.346
	6-311+G(2df,p)	1.906	2.333	1.892	2.344
	6-311++G(3df,3pd)	1.906	2.331	1.892	2.344
<b>ωB97XD</b>	6-31G(d)	1.908	2.310	1.892	2.333
	6-311G(d)	1.914	2.337	1.901	2.358
	6-311G(d,p)	1.913	2.338	1.900	2.358
	6-311+G(d,p)	1.913	2.338	1.898	2.357
	6-311G(2d,p)	1.910	2.339	1.897	2.361
	6-311G(df,p)	1.908	2.325	1.895	2.343
	6-311+G(2df,p)	1.906	2.324	1.893	2.341
	6-311++G(3df,3pd)	1.906	2.325	1.893	2.341
<b>M06-2X</b>	6-31G(d)	1.906	2.313	1.893	2.331
	6-311G(d)	1.914	2.336	1.903	2.356
	6-311G(d,p)	1.914	2.337	1.902	2.356
	6-311+G(d,p)	1.914	2.337	1.901	2.354
	6-311G(2d,p)	1.911	2.340	1.900	2.359
	6-311G(df,p)	1.909	2.326	1.897	2.343
	6-311+G(2df,p)	1.907	2.325	1.895	2.340
	6-311++G(3df,3pd)	1.907	2.326	1.895	2.341
<b>M08-HX</b>	6-31G(d)	1.910	2.306	1.892	2.326
	6-311G(d)	1.915	2.332	1.901	2.349
	6-311G(d,p)	1.915	2.333	1.901	2.349
	6-311+G(d,p)	1.914	2.333	1.904	2.372
	6-311G(2d,p)	1.911	2.335	1.897	2.353
	6-311G(df,p)	1.910	2.320	1.896	2.338
	6-311+G(2df,p)	1.908	2.321	1.893	2.334
	6-311++G(3df,3pd)	1.907	2.321	1.893	2.335
<b>QCISD</b>	6-311+G(2df,p)	1.914	2.329	1.902	2.359

**Appendix A3.4.** C–Se, C–S and S–Se bond lengths (ångström, Å) for CH<sub>3</sub>XYH and CH<sub>3</sub>XY<sup>−</sup> (X=S, Se; Y=Se, S).

Method	Basis set	CH <sub>3</sub> SSeH		CH <sub>3</sub> SSe <sup>−</sup>		CH <sub>3</sub> SeSH		CH <sub>3</sub> SeS <sup>−</sup>	
		C–S	S–Se	C–S	S–Se	C–Se	Se–S	C–Se	Se–S
<b>B3LYP</b>	6-31G(d)	1.837	2.219	1.836	2.266	1.973	2.226	1.985	2.234
	6-311G(d)	1.833	2.235	1.832	2.286	1.976	2.240	1.986	2.253
	6-311G(d,p)	1.834	2.237	1.833	2.286	1.975	2.242	1.986	2.253
	6-311+G(d,p)	1.834	2.237	1.833	2.281	1.975	2.241	1.988	2.244
	6-311G(2d,p)	1.832	2.227	1.832	2.270	1.973	2.231	1.983	2.244
	6-311G(df,p)	1.832	2.221	1.832	2.266	1.969	2.226	1.980	2.236
	6-311+G(2df,p)	1.826	2.213	1.826	2.250	1.968	2.216	1.981	2.221
	6-311++G(3df,3pd)	1.822	2.212	1.822	2.248	1.968	2.215	1.980	2.219
<b>B3PW91</b>	6-31G(d)	1.824	2.199	1.823	2.242	1.957	2.205	1.969	2.212
	6-311G(d)	1.820	2.212	1.820	2.260	1.960	2.217	1.970	2.228
	6-311G(d,p)	1.821	2.214	1.820	2.260	1.959	2.219	1.970	2.228
	6-311+G(d,p)	1.821	2.214	1.820	2.256	1.959	2.218	1.971	2.221
	6-311G(2d,p)	1.819	2.204	1.819	2.244	1.957	2.208	1.968	2.220
	6-311G(df,p)	1.819	2.199	1.819	2.241	1.953	2.204	1.964	2.212
	6-311+G(2df,p)	1.814	2.191	1.813	2.226	1.953	2.195	1.965	2.198
	6-311++G(3df,3pd)	1.810	2.190	1.809	2.224	1.953	2.194	1.964	2.196
<b>ωB97XD</b>	6-31G(d)	1.823	2.194	1.821	2.238	1.951	2.201	1.960	2.209
	6-311G(d)	1.820	2.207	1.818	2.256	1.947	2.190	1.962	2.224
	6-311G(d,p)	1.820	2.209	1.818	2.256	1.953	2.213	1.961	2.224
	6-311+G(d,p)	1.820	2.208	1.817	2.252	1.953	2.213	1.962	2.218
	6-311G(2d,p)	1.818	2.199	1.817	2.242	1.951	2.204	1.959	2.216
	6-311G(df,p)	1.818	2.195	1.817	2.239	1.948	2.200	1.957	2.209
	6-311+G(2df,p)	1.813	2.187	1.811	2.224	1.947	2.191	1.956	2.196
	6-311++G(3df,3pd)	1.809	2.186	1.807	2.223	1.947	2.190	1.956	2.195
<b>M06-2X</b>	6-31G(d)	1.822	2.193	1.820	2.236	1.950	2.199	1.958	2.210
	6-311G(d)	1.819	2.207	1.817	2.254	1.957	2.211	1.963	2.228
	6-311G(d,p)	1.819	2.208	1.818	2.254	1.956	2.212	1.963	2.228
	6-311+G(d,p)	1.819	2.208	1.817	2.250	1.956	2.212	1.964	2.220
	6-311G(2d,p)	1.817	2.200	1.816	2.239	1.954	2.203	1.960	2.221
	6-311G(df,p)	1.817	2.194	1.816	2.237	1.951	2.198	1.959	2.212
	6-311+G(2df,p)	1.812	2.188	1.811	2.221	1.950	2.190	1.959	2.198
	6-311++G(3df,3pd)	1.809	2.187	1.808	2.221	1.950	2.190	1.958	2.197
<b>M08-HX</b>	6-31G(d)	1.820	2.194	1.818	2.227	1.951	2.198	1.957	2.218
	6-311G(d)	1.818	2.207	1.816	2.245	1.957	2.209	1.962	2.232
	6-311G(d,p)	1.819	2.208	1.816	2.245	1.956	2.210	1.962	2.232
	6-311+G(d,p)	1.818	2.208	1.816	2.240	1.956	2.210	1.964	2.224
	6-311G(2d,p)	1.816	2.200	1.814	2.229	1.954	2.201	1.959	2.225
	6-311G(df,p)	1.817	2.195	1.815	2.228	1.951	2.197	1.958	2.217
	6-311+G(2df,p)	1.811	2.188	1.809	2.213	1.950	2.190	1.958	2.203
	6-311++G(3df,3pd)	1.809	2.187	1.807	2.212	1.951	2.189	1.958	2.201
<b>QCISD</b>	6-311+G(2df,p)	1.819	2.198	1.816	2.239	1.952	2.202	1.961	2.207

**Appendix A3.5.** Optimized S–S bond lengths (ångström, Å) for RSSSH/<sup>−</sup> (R=CH<sub>3</sub>, CH<sub>2</sub>=CH).

Method	Basis set	CH <sub>3</sub> SSSH		CH <sub>3</sub> SSS <sup>−</sup>		CH <sub>2</sub> CHSSSH		CH <sub>2</sub> CHSSS <sup>−</sup>	
		S <sub>1</sub> –S <sub>2</sub>	S <sub>2</sub> –S <sub>3</sub>	S <sub>1</sub> –S <sub>2</sub>	S <sub>2</sub> –S <sub>3</sub>	S <sub>1</sub> –S <sub>2</sub>	S <sub>2</sub> –S <sub>3</sub>	S <sub>1</sub> –S <sub>2</sub>	S <sub>2</sub> –S <sub>3</sub>
<b>B3LYP</b>	6-31G(d)	2.079	2.113	2.176	2.073	2.095	2.105	2.206	2.059
	6-311G(d)	2.089	2.123	2.183	2.088	2.108	2.114	2.213	2.071
	6-311+G(d)	2.088	2.122	2.176	2.085	2.104	2.113	2.206	2.069
	6-311+G(d,p)	2.088	2.124	2.176	2.085	2.103	2.115	2.206	2.068
	6-311G(2d,p)	2.075	2.108	2.147	2.077	2.091	2.101	2.169	2.064
	6-311+G(df,p)	2.076	2.112	2.152	2.077	2.089	2.104	2.179	2.062
	6-311+G(2df,p)	2.060	2.094	2.122	2.061	2.075	2.086	2.147	2.047
	6-311++G(3df,3pd)	2.052	2.087	2.110	2.053	2.066	2.080	2.135	2.040
<b>B3PW91</b>	6-31G(d)	2.062	2.092	2.145	2.057	2.076	2.085	2.169	2.044
	6-311G(d)	2.070	2.100	2.151	2.070	2.086	2.091	2.174	2.055
	6-311+G(d)	2.070	2.099	2.147	2.067	2.083	2.091	2.171	2.052
	6-311+G(d,p)	2.069	2.102	2.147	2.067	2.082	2.094	2.170	2.052
	6-311G(2d,p)	2.057	2.086	2.120	2.058	2.071	2.079	2.136	2.047
	6-311+G(df,p)	2.058	2.090	2.126	2.059	2.070	2.084	2.145	2.046
	6-311+G(2df,p)	2.044	2.073	2.097	2.043	2.056	2.066	2.117	2.032
	6-311++G(3df,3pd)	2.036	2.067	2.087	2.036	2.048	2.061	2.105	2.025
<b>ωB97XD</b>	6-31G(d)	2.061	2.084	2.122	2.063	2.073	2.077	2.136	2.053
	6-311G(d)	2.069	2.091	2.127	2.077	2.084	2.084	2.139	2.067
	6-311+G(d)	2.069	2.091	2.124	2.074	2.080	2.084	2.137	2.063
	6-311+G(d,p)	2.068	2.093	2.124	2.073	2.078	2.086	2.134	2.062
	6-311G(2d,p)	2.057	2.080	2.105	2.063	2.070	2.074	2.115	2.054
	6-311+G(df,p)	2.058	2.083	2.108	2.065	2.067	2.077	2.118	2.055
	6-311+G(2df,p)	2.044	2.068	2.087	2.048	2.054	2.062	2.098	2.039
	6-311++G(3df,3pd)	2.037	2.062	2.078	2.042	2.047	2.056	2.089	2.032
<b>M06-2X</b>	6-31G(d)	2.060	2.082	2.121	2.063	2.072	2.076	2.144	2.053
	6-311G(d)	2.067	2.088	2.124	2.078	2.083	2.081	2.144	2.067
	6-311+G(d)	2.067	2.088	2.121	2.074	2.080	2.082	2.143	2.063
	6-311+G(d,p)	2.066	2.090	2.121	2.074	2.079	2.083	2.143	2.063
	6-311G(2d,p)	2.056	2.078	2.103	2.065	2.070	2.071	2.118	2.055
	6-311+G(df,p)	2.055	2.079	2.104	2.065	2.067	2.074	2.119	2.055
	6-311+G(2df,p)	2.043	2.065	2.085	2.049	2.056	2.059	2.100	2.039
	6-311++G(3df,3pd)	2.036	2.060	2.077	2.042	2.048	2.055	2.092	2.032
<b>M08-HX</b>	6-31G(d)	2.063	2.080	2.124	2.063	2.081	2.071	2.144	2.053
	6-311G(d)	2.071	2.088	2.128	2.076	2.089	2.078	2.148	2.065
	6-311+G(d)	2.070	2.087	2.126	2.073	2.085	2.079	2.148	2.061
	6-311+G(d,p)	2.070	2.089	2.126	2.073	2.084	2.081	2.149	2.061
	6-311G(2d,p)	2.060	2.077	2.110	2.064	2.075	2.069	2.125	2.055
	6-311+G(df,p)	2.059	2.079	2.110	2.064	2.072	2.071	2.130	2.053
	6-311+G(2df,p)	2.046	2.064	2.090	2.049	2.060	2.057	2.108	2.039
	6-311++G(3df,3pd)	2.041	2.059	2.083	2.041	2.053	2.053	2.101	2.032
<b>QCISD</b>	6-311+G(2df,p)	2.056	2.078	2.096	2.065	2.107	2.057	2.066	2.073

## Appendix A: Polysulfur/Selenium Assessment

### Appendix A3.6. Mulliken charges on every S of RSSSH<sup>-</sup> (R=CH<sub>3</sub>, CH<sub>2</sub>=CH).

Method/Basis Set	CH <sub>3</sub> SSSH			CH <sub>3</sub> SSS <sup>-</sup>		
	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>
QCISD/6-311+G(2df,p)	0.15	0.05	-0.25	-0.05	-0.10	-0.77
Method/Basis Set	CH <sub>2</sub> CHSSSH			CH <sub>2</sub> CHSSS <sup>-</sup>		
	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>
QCISD/6-311+G(2df,p)	-0.01	0.04	-0.20	-0.23	-0.08	-0.74
Method/Basis Set	CYS-SSSH			CYS-SSS <sup>-</sup>		
	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>
ωB97XD/6-311G(2d,p)	0.00	-0.03	-0.12	-0.12	-0.13	-0.66

**Appendix A3.7.** Homolytic S–S bond dissociation enthalpy (BDE) of RSSH, proton affinity (PA) and gas-phase basicity (GPB) of RSS<sup>-</sup>, and hydrogen affinity (HA) of RSS<sup>•</sup> (R=CH<sub>3</sub>, CH<sub>2</sub>CH). All energies calculated at 298.15K and in kJ mol<sup>-1</sup>.

Methods	Basis set	CH <sub>3</sub> SSH	CH <sub>3</sub> SS <sup>-</sup>	CH <sub>3</sub> SS <sup>•</sup>	CH <sub>2</sub> CHSSH	CH <sub>2</sub> CHSS <sup>-</sup>	CH <sub>2</sub> CHSS <sup>•</sup>		
		BDE(S—S)	PA	GPB	HA	BDE(S—S)	PA	GPB	HA
B3LYP	6-31G(d)	224.8	1456.4	1426.1	276.4	228.1	1430.0	1398.6	274.7
	6-311G(2d,p)	235.6	1440.5	1410.1	285.6	239.2	1418.0	1386.0	259.6
	6-311G+(2df,p)	240.3	1437.7	1407.3	281.1	241.9	1411.5	1380.2	280.2
	6-311G++(3df,3pd)	241.6	1440.9	1410.6	283.5	243.2	1414.7	1383.5	282.1
B3PW91	6-31G(d)	235.4	1455.2	1424.8	274.6	238.0	1427.6	1396.5	272.8
	6-311G(2d,p)	250.3	1447.7	1417.3	283.8	253.0	1423.1	1391.5	254.5
	6-311G+(2df,p)	254.7	1445.8	1415.4	279.5	255.7	1417.7	1386.8	278.4
	6-311G++(3df,3pd)	255.8	1448.7	1418.4	281.6	256.8	1420.5	1389.7	280.1
ωB97XD	6-31G(d)	244.1	1455.0	1425.1	284.0	244.1	1429.8	1397.3	283.7
	6-311G(2d,p)	258.5	1447.3	1417.5	292.0	258.5	1422.4	1395.5	291.5
	6-311G+(2df,p)	264.6	1445.5	1415.8	288.0	262.7	1420.2	1382.0	288.3
	6-311G++(3df,3pd)	265.6	1449.2	1419.6	289.8	263.5	1423.6	1388.6	289.7
M06-2X	6-31G(d)	252.0	1440.0	1411.4	285.4	249.8	1414.5	1381.1	283.7
	6-311G(2d,p)	264.9	1430.0	1401.6	294.4	263.0	1408.2	1374.4	292.8
	6-311G+(2df,p)	270.3	1428.5	1400.1	291.0	267.4	1403.4	1366.4	290.2
	6-311G++(3df,3pd)	272.3	1432.3	1404.2	293.9	269.1	1406.7	1368.1	292.6
M08-HX	6-31G(d)	258.0	1446.1	1414.2	295.4	255.5	1419.6	1387.9	291.6
	6-311G(2d,p)	268.3	1435.5	1405.6	298.2	268.7	1414.3	1382.0	297.1
	6-311G+(2df,p)	273.6	1433.5	1404.3	294.0	274.0	1411.4	1376.4	296.2
	6-311G++(3df,3pd)	275.2	1437.7	1408.6	296.4	272.7	1412.9	1383.4	295.7
QCISD	6-311G+(2df,p)	236.3	1445.4	1415.1	282.7	242.0	1421.8	1390.7	283.5

**Appendix A3.8.** Homolytic Se–Se bond dissociation enthalpy (BDE) of RSeSeH, proton affinity (PA) and gas-phase basicity (GPB) of RSeSe<sup>−</sup>, and hydrogen affinity (HA) of RSeSe<sup>•</sup> (R=CH<sub>3</sub>, CH<sub>2</sub>CH). All energies calculated at 298.15K and in kJ mol<sup>−1</sup>.

Methods	Basis set	CH <sub>3</sub> SeSeH	CH <sub>3</sub> SeSe <sup>−</sup>		CH <sub>3</sub> SeSe <sup>•</sup>	CH <sub>2</sub> CHSeSeH	CH <sub>2</sub> CHSeSe <sup>−</sup>		CH <sub>2</sub> CHSeSe <sup>•</sup>
		BDE(Se—Se)	PA	GPB	HA	BDE(Se—Se)	PA	GPB	HA
<b>B3LYP</b>	6-311+G(d,p)	208.0	1401.6	1371.3	272.2	211.5	1382.6	1350.9	272.0
	6-311G(df,p)	212.9	1421.0	1390.7	267.0	193.2	1400.0	1369.1	266.2
	6-311G+(2df,p)	212.0	1404.0	1373.7	267.7	214.9	1384.8	1353.2	267.1
	6-311G++(3df,3pd)	212.7	1404.3	1374.1	268.5	215.4	1385.5	1353.9	268.0
<b>B3PW91</b>	6-311+G(d,p)	219.2	1409.0	1378.7	270.5	222.1	1388.6	1357.4	270.2
	6-311G(df,p)	223.7	1424.2	1393.8	265.3	204.7	1401.9	1371.0	264.5
	6-311G+(2df,p)	222.9	1411.4	1381.1	266.0	225.3	1390.7	1359.7	265.4
	6-311G++(3df,3pd)	223.6	1411.8	1381.5	266.7	226.6	1392.1	1360.7	266.7
<b>ωB97XD</b>	6-311+G(d,p)	218.0	1410.8	1380.7	276.5	221.3	1391.4	1359.6	277.1
	6-311G(df,p)	223.7	1426.5	1396.4	270.3	226.2	1405.4	1373.9	270.5
	6-311G+(2df,p)	222.6	1413.3	1383.4	270.7	225.3	1393.8	1361.1	271.0
	6-311G++(3df,3pd)	223.7	1413.6	1383.7	271.4	226.3	1391.9	1366.1	271.8
<b>M06-2X</b>	6-311+G(d,p)	203.5	1384.2	1353.6	275.3	206.2	1367.6	1335.5	277.3
	6-311G(df,p)	209.0	1397.8	1367.0	271.6	212.0	1377.0	1352.2	273.0
	6-311G+(2df,p)	207.7	1386.0	1355.3	272.2	211.0	1369.0	1336.1	273.6
	6-311G++(3df,3pd)	223.4	1387.4	1356.8	273.4	212.4	1370.5	1337.8	272.3
<b>M08-HX</b>	6-311+G(d,p)	233.7	1380.0	1349.7	291.4	235.9	*	*	291.8
	6-311G(df,p)	237.6	1393.8	1363.4	288.1	240.7	1373.8	1339.7	288.5
	6-311G+(2df,p)	223.4	1381.9	1351.6	288.8	224.9	1362.6	1329.0	288.5
	6-311G++(3df,3pd)	223.4	1383.4	1353.2	289.5	225.7	1362.8	1337.3	290.0
<b>QCISD</b>	6-311G+(2df,p)	204.1	1404.3	1374.0	267.3	208.3	1386.7	1355.7	265.9

\* Poor optimization of anion led to high energy conformation.

**Appendix A3.9.** Homolytic X–Y bond dissociation enthalpy (BDE) of RXYH, proton affinity (PA) and gas-phase basicity (GPB) of RXY<sup>−</sup>, and hydrogen affinity (HA) of RXY<sup>•</sup> (R=CH<sub>3</sub>, CH<sub>2</sub>CH; X=S, Se; Y=Se, S). All energies calculated at 298.15K and in kJ mol<sup>−1</sup>.

		CH <sub>3</sub> SSeH	CH <sub>3</sub> SSe <sup>−</sup>	CH <sub>3</sub> SSe <sup>•</sup>		CH <sub>3</sub> SeSH	CH <sub>3</sub> SeS <sup>−</sup>	CH <sub>3</sub> SeS <sup>•</sup>	
Methods	Basis set	BDE(S—Se)	PA	GPB	HA	BDE(Se—S)	PA	GPB	HA
B3LYP	6-311+G(d,p)	209.8	1405.0	1374.8	275.2	217.3	1427.8	1397.6	289.2
	6-311G(df,p)	215.2	1421.2	1391.0	269.8	223.4	1435.1	1404.8	283.5
	6-311G+(2df,p)	220.7	1407.8	1377.7	267.0	229.0	1434.1	1403.9	285.3
	6-311G++(3df,3pd)	220.7	1408.0	1377.9	267.6	229.1	1437.2	1407.0	288.6
B3PW91	6-311+G(d,p)	221.7	1413.2	1382.8	273.8	229.4	1435.0	1404.8	287.1
	6-311G(df,p)	227.1	1425.8	1395.7	268.2	235.2	1440.5	1410.1	281.7
	6-311G+(2df,p)	233.3	1415.7	1385.5	265.3	241.6	1441.7	1411.4	283.6
	6-311G++(3df,3pd)	233.2	1415.9	1385.8	265.9	241.5	1444.5	1414.3	286.5
ωB97XD	6-311+G(d,p)	219.6	1414.0	1384.2	279.9	231.4	1435.6	1405.8	295.6
	6-311G(df,p)	226.1	1427.3	1397.5	273.9	238.5	1440.7	1410.9	289.3
	6-311G+(2df,p)	232.5	1417.2	1387.5	270.9	245.0	1442.2	1412.4	290.8
	6-311G++(3df,3pd)	232.3	1417.6	1388.0	271.6	244.9	1445.2	1415.7	293.6
M06-2X	6-311+G(d,p)	222.3	1390.1	1359.8	277.7	226.8	1415.5	1384.6	297.8
	6-311G(df,p)	229.2	1400.9	1370.7	268.1	234.5	1420.3	1389.4	294.9
	6-311G+(2df,p)	234.3	1392.1	1362.0	270.5	240.0	1422.3	1391.7	296.3
	6-311G++(3df,3pd)	235.0	1393.2	1363.2	271.9	254.9	1425.7	1395.1	299.5
M08-HX	6-311+G(d,p)	231.7	1386.7	1357.9	293.8	242.0	1420.6	1390.1	309.5
	6-311G(df,p)	236.8	1397.4	1368.2	288.3	246.5	1426.9	1396.1	306.6
	6-311G+(2df,p)	228.4	1387.9	1359.1	285.2	203.1	1382.7	1260.1	263.3
	6-311G++(3df,3pd)	227.9	1388.9	1359.7	285.8	259.9	1431.8	1401.2	311.8
QCISD	6-311G+(2df,p)	218.7	1411.3	1381.1	267.5	224.5	1438.9	1408.6	286.1

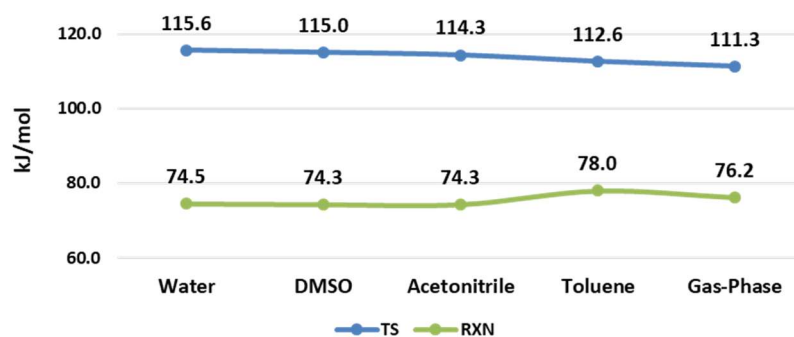
**Appendix A3.10.** Homolytic S–S bond dissociation enthalpy (BDE) of RSSSH, proton affinity (PA) and gas-phase basicity (GPB) of  $\text{RSSS}^-$ , and hydrogen affinity (HA) of  $\text{RSSS}^\bullet$  ( $\text{R}=\text{CH}_3, \text{CH}_2\text{CH}$ ). All energies calculated at 298.15K and in  $\text{kJ mol}^{-1}$ .

		CH <sub>3</sub> SSSH	CH <sub>3</sub> SSS <sup>-</sup>	CH <sub>3</sub> SSS <sup>•</sup>		CH <sub>2</sub> CHSSSH	CH <sub>2</sub> CHSSS <sup>-</sup>	CH <sub>2</sub> CHSSS <sup>•</sup>	
Methods	Basis set	BDE(SS—S)	PA	GPB	HA	BDE(SS—S)	PA	GPB	HA
B3LYP	6-31G(d)	167.6	1415.2	1383.7	283.9	166.9	1402.9	1369.7	288.1
	6-311G(2d,p)	178.4	1407.1	1375.6	292.8	177.7	1384.9	1351.4	297.2
	6-311G+(2df,p)	180.4	1406.7	1375.2	289.5	179.7	1392.7	1360.0	292.9
	6-311G++(3df,3pd)	180.1	1410.1	1378.7	291.9	179.0	1396.1	1363.4	294.9
B3PW91	6-31G(d)	177.9	1413.9	1382.6	282.6	176.8	1427.5	1394.2	287.1
	6-311G(2d,p)	191.5	1413.9	1382.3	290.9	190.5	1401.8	1369.3	293.9
	6-311G+(2df,p)	193.8	1413.9	1382.4	287.9	192.8	1399.5	1367.0	290.5
	6-311G++(3df,3pd)	195.9	1419.4	1387.9	292.3	192.0	1402.4	1370.1	292.2
ωB97XD	6-31G(d)	184.2	1415.6	1385.2	293.1	185.0	1406.0	1376.3	297.8
	6-311G(2d,p)	197.3	1415.1	1385.0	300.8	198.0	1405.6	1375.1	305.5
	6-311G+(2df,p)	200.4	1415.0	1385.1	297.7	201.0	1403.0	1370.9	302.0
	6-311G++(3df,3pd)	199.6	1418.7	1388.8	299.7	199.4	1406.3	1375.3	303.8
M06-2X	6-31G(d)	192.6	1398.7	1369.0	292.5	192.7	1389.5	1357.5	297.3
	6-311G(2d,p)	204.7	1395.6	1365.4	300.9	205.2	1387.3	1356.3	306.5
	6-311G+(2df,p)	208.6	1395.7	1365.3	298.4	208.7	1385.4	1356.1	303.2
	6-311G++(3df,3pd)	209.1	1399.2	1368.7	300.9	208.6	1388.6	1358.7	305.7
M08-HX	6-31G(d)	196.0	1402.8	1376.5	300.8	195.7	1393.8	1362.2	305.8
	6-311G(2d,p)	206.8	1400.7	1371.2	304.5	206.9	1392.9	1361.8	310.1
	6-311G+(2df,p)	209.6	1400.6	1371.4	301.3	209.4	1389.8	1359.3	306.1
	6-311G++(3df,3pd)	209.8	1404.8	1376.2	303.7	208.8	1393.8	1363.8	308.3
QCISD	6-311G+(2df,p)	183.0	1413.9	1384.1	289.4	185.7	1403.7	1371.2	295.0

## **Appendix B**



**Appendix B4.11.** Calculated Gibbs's Free Energy activation barriers (difference in energy between reactants and transition structure; blue line) and Gibbs's Free Energis of reactions (difference in energy between reactants and products; green line) for the reaction of SO<sub>2</sub> with N-tosyl hydrazone in various solvents and in the gas-phase. It is noted that the solvents are shown in order of decreasing dielectric constant.



## VITA AUCTORIS

NAME: Sahar Nikoo

PLACE OF BIRTH: Tehran, Iran

YEAR OF BIRTH: 1981

EDUCATION: Ph.D., Chemistry and Biochemistry

University of Windsor

Windsor, ON, Canada, 2019

M.Sc., Organic Chemistry

Islamic Azad University

Tehran, Iran, 2008

B.Sc., Applied Chemistry

Islamic Azad University

Tehran, Iran, 2005